مقدمة في التحليل الإحصائي

لبحوث الإعلام

أ. د. بركات عبد العزيز

عبد العزيز ، بركات.
مقدمة في التحليل الإحصائي لبحوث الإعلام / بركات عبد العزيز
. ــ ط١. ــ القاهرة: الدار المصرية اللبنانية، 2014
424 ص؛24 سم.
1 ــ الإحصاء التحليلي
٢- الإعلام
أ ــ العنوان 519.5
رقم الإيداع: 5523 / 2014

الدار المصرية اللبنانية
16 عبد الخالق ثروت القاهرة.
تليفون: 23910250 ٢٠٢ +
فاكس: 23909618 ٢٠٢ + ــ ص.ب 2022
E-mail:info@almasriah.com
www.almasriah.com

رئيس مجلس الإدارة: محمد رشـــاد
المشرف الفني: محمد حجي

المكتبة الإعلامية

الطبعـة الأولى: جماد أول 1435هــ ــ مارس ٢٠١٤م

مقدمة في التحليل الإحصائي لبحوث الإعلام

أ.د. بركات عبد العزيز
كلية الإعلام - جامعة القاهرة

الدار المصرية اللبنانية

بسم اللـه الرحمن الرحيم

المكتبـة الإعـلاميـة

من منطلق حرص الدار المصرية اللبنانية على إصدار سلاسل متخصصة في مختلف العلوم والفنون والآداب، تأتي هذه السلسلة (المكتبة الإعلامية) لتتكامل مع سلاسل أخرى أصدرتها الدار في العلوم التربوية والدينية والأدبية والفكرية، بما يسمح بسهولة متابعة الإنتاج الفكري الجديد لكافة الدارسين والممارسين.

وتهدف هذه السلسلة إلى تحقيق الأغراض التالية:

١ - إثراء المكتبة العربية في مجالات علوم الاتصال وفنون الإعلام؛ حيث شهدت هذه العلوم تطورات كبيرة خلال القرن العشرين، وأصبح الإعلام ظاهرة مؤثرة في جميع الأنشطة السياسية والاقتصادية والاجتماعية.

٢ - ظهور عديد من كليات وأقسام الإعلام في الجامعات المصرية والعربية، وحاجة هذه الأقسام إلى متابعة الإنتاج الفكري في مجال الإعلام، الذي يسهم في تنظير فروع علم الاتصال من منظور عربي.

٣ - تزويد الممارسين للعمل الإعلامي بالمعلومات الجديدة في مجالات التكنولوجيا والإنتاج الإعلامي، وتأثير الرسائل الإعلامية والإعلانية على الجماهير المستهدفة.

٤ - نشر الثقافة الإعلامية من خلال التأليف والترجمة ونشر الرسائل المتميزة للماجستير والدكتوراه؛ لأهمية هذه الثقافة التي أصبحت ضرورة لا غنى عنها، لتيسير الانتفاع بمصادر المعلومات والإعلام المتعددة في العصر الحديث.

الناشر

المحتويات

الفصل الثاني

إعداد البيانات والمعالجة الإحصائية الوصفية

الفصل الثالث

مقاييس العلاقة والارتباط

الفصل الرابع

مقاييس الفروق

الفصل الخامس

تعريف ببعض الأساليب الإحصائية المتعمقة

الفصل السادس

العينات في البحث العلمي

***

تقديـــم

يمثل التحليل الإحصائي أهمية قصوى للتثبت من دقة نتائج البحوث العلمية في كافة التخصصات، ولا تكاد تخلو دراسة علمية رصينة من دلالات التحليل الإحصائي، خاصة فيما يتعلق ببحوث التسويق، واستطلاعات الرأي العام، ورسائل الماجستير والدكتوراه، وبحوث الترقيات العلمية للحصول على لقب أستاذ مساعد وأستاذ.

وتبرز أهمية التحليل الإحصائي في بحوث الإعلام بكافة مجالاته ومستوياته، وتجدر الإشارة إلى معاناة الكثير من الباحثين من عدم القدرة على توظيف علم الإحصاء وتطبيقاته على بحوث الإعلام، وبالتالي يلجأون إلى الاستعانة بأشخاص آخرين لتنفيذ الجانب الإحصائي على نتائج بحوثهم الكميّة، وقد يكون بعض هؤلاء من غير الدارسين للإعلام، وبالتالي يستخدمون معالجات إحصائية قد لا تخدم البحث العلمي ولا تساعد على تفسير النتائج.

ومن هنا تأتي أهمية هذا الكتاب الذي يطرح أساليب استخدام التحليل الإحصائي في بحوث الإعلام، بأسلوب بسيط وواضح يناسب الطالب والباحث في العلوم التطبيقية للإعلام.

يطرح الكتاب خلفية معرفية حول أهمية علم الإحصاء لنشر البحوث العلمية، والفرق بين الإحصاء الوصفي والإحصاء الاستدلالي، واستخدام الحاسب الآلي في التحليل الإحصائي، مع التركيز على الحزمة الإحصائية للعلوم الاجتماعية (SPSS). كما يشرح الكتاب كيفية إدخال البيانات وإعدادها للتحليل الإحصائي، وعرض لمقاييس العلاقة والارتباط، ومقاييس الفروق بين المجموعات من حيث النسب أو المتوسطات، والأساليب الأكثر تعمقًا مثل: تحليل التباين المتعدد، وتحليل التغاير، وتحليل الانحدار،

والتحليل العاملي، وتحليل المسار، والتحليل التمييزي.. باعتبارها من أكثر الأساليب استخدامًا في العلوم الاجتماعية، فضلًا عن تخصيص فصل كامل للعينات وأسلوب اختيارها، والطرق الإحصائية لتقدير حجم العينة.

ويبقى أن نؤكد على أن مؤلف الكتاب الدكتور بركات عبد العزيز أحد الأساتذة البارزين في مناهج البحث والرأي العام والإحصاء، ولديه خبرات عميقة في هذا المجال على المستويين النظري والتطبيقي؛ مما يجعل من هذا الكتاب إضافة علمية متميزة للمكتبة الإعلامية.

هيئة التحرير

المقدمـــة

الإحصاء هو الأساس القاعدي للبحث العلمي في كافة فروع المعرفة، وقد أصبحت بحوث الاتصال تعتمد على التحليل الإحصائي بكثافة ملحوظة؛ الأمر الذي ساعد كثيرًا على تطور هذه البحوث واتساع نطاقها، وكان لذلك أثره في تزايد حاجة الباحثين والدارسين إلى المعرفة والمهارات الإحصائية الحتي تناسب «خصوصية» موضوعات الاتصال في أدواته التقليدية والحديثة. من هنا تم إعداد هذه المادة العلمية المختصرة كمبادئ أساسية عامة للتحليل الإحصائي، وذلك من خلال ستة فصول، يتضمن الفصل الأول مدخلًا تعريفيًا بالإحصاء ودلالته للبحث العلمي في جوانبه المتطورة. ويضم هذا الفصل ثلاثة مباحث، يتناول المبحث الأول تعريفًا موجزًا بعلم الإحصاء من منظور أهميته للبحث العلمي، ومن ذلك نوضح مسألة القياس الإحصائي مع التركيز على المتغيرات بما في ذلك توضيح الفرق بين الإحصاء الوصفي والإحصاء الاستدلالي، وكذلك الإحصاء البارامتري والإحصاء اللابارامتري. وفي المبحث الثاني توضيح لأهم الرموز والعمليات الجبرية الإحصائية، أما المبحث الثالث فيتضمن نبذة مختصرة عن استخدام الحاسب الآلي في التحليل الإحصائي من خلال تعريف موجز بمزايا الحاسوب وحدوده والبرامج الإحصائية الجاهزة مع التركيز على الحزمة الإحصائية للعلوم الاجتماعية (SPSS).

أما الفصل الثاني فهو عن كيفية إعداد البيانات والمعالجة الإحصائية الوصفية؛ حيث يشرح الفصل أساسيات إدخال البيانات وإعدادها للتحليل الإحصائي من منظور الفكر والمهارة، وذلك عبر ثلاثة مباحث، يختص المبحث الأول بتوضيح عملية تجهيز البيانات للتحليل الإحصائي، بمعنى كيفية إعدادها بطريقة صحيحة حتى يمكن معالجتها إحصائيًا

معالجة سليمة تقابل أهداف البحث؛ وذلك من حيث: المراجعة، الإدخال، التنظيم. أما المبحث الثاني فيناقش أساليب الوصف التكراري والبياني للمعطيات الإحصائية الأولية من خلال التكرارات والنسب والأشكال البيانية، في حين يأتي المبحث الثالث متضمنا التعريف بمقاييس النزعة المركزية للبيانات، ومقاييس التشتت.

وفي الفصل الثالث توضيح مختصر لمقاييس العلاقة والارتباط باعتبار أن هذه المقاييس تقوم بدور أساسي في مختلف البحوث العلمية، ولا يكاد يخلو منها أي بحث إمبيريقي سواء بشكل مباشر أو غير مباشر، بل إن معظم الأساليب الإحصائية تقوم على معاملات الارتباط، وقد لوحظ أن هناك بعض المثالب في اختيار وتطبيق معاملات الارتباط خاصة من حيث اختيار النمط الذي يناسب المتغيرات المطلوب تقصي الارتباط بينها. وعلى هذا الأساس فإن الفصل الثالث يلقي الضوء على الجوانب الأساسية لمقاييس العلاقة والارتباط من خلال توضيح معنى الارتباط وخصائصه، والاختيار الصحيح لمعاملات الارتباط، وأهم معاملات الارتباط شائعة الاستخدام في البحوث الاجتماعية مع الإشارة إلى الارتباط الجزئي، وكذلك معامل الارتباط المتعدد، باعتبارهما من أدوات الضبط الإحصائي المهمة سواء في الدراسات المسحية أو التجريبية.

وفي سياق متكامل يأتي الفصل الرابع ليشرح مقاييس الفروق؛ حيث يتضمن تعريفًا مختصرًا ببعض الأساليب الإحصائية التي تقيس الفروق بين المجموعات المختلفة، سواء من حيث النسب (Percentages) أو من حيث المتوسطات (Means)، ويضم هذا الفصل ثلاثة مباحث، يتناول المبحث الأول أساليب قياس الفروق بين مجموعتين، سواء فيما يخص النسب، أو فيما يخص المتوسطات. أما المبحث الثاني فيتناول أساليب قياس الفروق بين أكثر من مجموعتين من حيث متوسطات الدرجات على مقياس معين، مع الأخذ في الاعتبار تعدد المتغيرات وتفاعلها، في حين يتناول المبحث الثالث أهم الأساليب الإحصائية اللابارامترية في قياس الفروق بين المجموعات.

أما الفصل الخامس فهو مدخل تعريفي ببعض الأساليب الإحصائية المتعمقة،ذلك أن الأساليب الإحصائية تتطور باستمرار، مستفيدة من التطور المتلاحق للبرامج الإحصائية الجاهزة وتطبيقاتها في كافة مجالات المعرفة، وعلى الرغم من أن الكثير من الأساليب الإحصائية المتعمقة حيز الاستخدام منذ قرابة منتصف القرن العشرين، إلا أن التطورات التي يتم إدخالها على تلك الأساليب تجعلها ذات أبعاد نظرية وتطبيقية أوسع نطاقًا وأكثر جدوى للبحوث العلمية، وبناءً على ذلك فإن الفصل الخامس يتضمن أمثلة لتلك الأساليب ممثلة في: تحليل التباين المتعدد، تحليل التغاير، تحليل الانحدار، التحليل العاملي،تحليل المسار، التحليل التمييزي.. باعتبارها من أكثر الأساليب استخدامًا في العلوم الاجتماعية، مع الإشارة إلى التحليل الثانوي والتحليل البعدي.

أخيرًا، فإن الفصل السادس من هذا الكتاب يتناول العينات في البحث العلمي، وذلك من خلال ثلاثة مباحث، يتضمن المبحث الأول مفهوم وحجم العينة بما في ذلك أهم الطرق الإحصائية لتقدير حجم العينة، أما المبحث الثاني فيتناول أنواع العينات (العينات العشوائية والعينات غير العشوائية)، ثم يأتي المبحث الثالث موضحًا أخطاء المعاينة ممثلة في الخطأ المعياري للمتوسط، والخطأ المعياري للنسبة، والخطأ المعياري لمعامل الارتباط مع توضيح كيفية الاستفادة من نتائج العينات في تقدير الظاهرة في المجتمع عند درجات ثقة مختلفة.

هذه المادة العلمية روعي فيها التركيز والتبسيط؛ بحيث تكون سهلة الفهم والتطبيق، وتستثير أذهان الدارسين والباحثين نحو المزيد من التعمق في التحليل الإحصائي؛ لامتلاك المعلومات الأساسية ومهارات التطبيق، وتكوين الاتجاهات الإيجابية نحو علم الإحصاء الذي كان له الفضل في تطور البحث العلمي في كافة مجالات المعرفة بما فيها مجال الإعلام والاتصال بالجماهير.

المؤلف

أ.د/ بركات عبد العزيز

الفصل الأول

مدخل تمهيدي

تمهـــيد

يتضمن هذا الفصل مدخلًا تعريفيًا بالإحصاء ودلالته للبحث العلمي في جوانبه المتطورة، وذلك من خلال ثلاثة مباحث. يتناول المبحث الأول تعريفًا موجزًا بعلم الإحصاء من منظور أهميته للبحث العلمي، ومن ذلك نوضح مسألة القياس الإحصائي مع التركيز على المتغيرات وتوضيح الفرق بين الإحصاء الوصفي والإحصاء الاستدلالي، وكذلك الإحصاء البارامتري والإحصاء اللابارامتري. ويتناول المبحث الثاني أهم الرموز والعمليات الجبرية الإحصائية. أما المبحث الثالث فهو عن استخدام الحاسب الآلي في التحليل الإحصائي من خلال تعريف موجز بمزايا الحاسوب وحدوده والبرامج الإحصائية الجاهزة بالتطبيق على الحزمة الإحصائية للعلوم الاجتماعية (SPSS)، باعتبارها الأكثر انتشارًا من حيث الاستخدام في مصر والمنطقة العربية.

المبحث الأول

الإحصاء من منظور أهميته للبحث العلمي

أولاً: الإحصاء وأهميته للعلم

الإحصاء في اللغة يعني العد الشامل، ومن المجاز قول العرب: «لم أر أكثر منهم حصى»؛ أي لم أر أكثر منهم عددًا، وقولهم: «هذا أمر لا أحصيه»؛ أي لا أطيقه ولا أضبطه، وقد وردت كلمة «الإحصاء» ومشتقاتها في الكثير من المواضع بالقرآن

الكريم. أما الإحصاء كممارسة فقد نشأ في إطار التنظيم السياسي للدولة القومية في أوروبا منذ الثلث الأخير من القرن الثامن عشر (عام ١٧٧٠)، غير أنه سرعان ما تطورت الأصول الرياضية لعلم الإحصاء بفضل علماء الرياضيات الأوروبيين وغيرهم، والذين تواصلت جهودهم العلمية حتى وصل الإحصاء إلى ما هو عليه الآن كعلم متطور يرتبط بجميع مجالات الحياة.

وإذا كان الإحصاء أحد فروع الرياضيات التي تطبق في مجالات البحث العلمي بفروع المعرفة كافة، فإنه بصورته الحديثة دعامة أساسية لطرق البحث العلمي ذات الطبيعة الكميّة (Quantitative)، والتي تقوم على إجراء دراسات ميدانية وتحليلية، وتجارب وملاحظات ومقابلات، والحصول منها على بيانات يتم تنظيم تلك البيانات ومعالجتها؛ بحيث يتم استخلاص نتائج موضوعية تصف الموضوع أو الظاهرة، وقد تتعدى نطاق الوصف إلى التفسير وصياغة القوانين والنظريات العلمية. ومن خلال التحليل الإحصائي لبيانات البحوث يتم التعرف على الدلالات الكمية للظواهر المختلفة في أبعادها الوصفية والتحليلية والتفسيرية والتوقعية؛ أي أن الإحصاء ليس مجرد عرض بيانات رقمية في جداول أو أشكال بيانية ورسوم هندسية، وإنما يتعدى ذلك إلى التحليل المتعمق لتلك البيانات وتفسيرها واستنطاق ما وراءها وبناء نماذج تفسيرية وصياغة توقعات مستقبلية عن الظواهر والموضوعات مجال البحث بما يفيد في الوصف والتفسير والتنبؤ واتخاذ القرارات.

وبموجب المعالجة الإحصائية السليمة يمكن رصد النتائج رصدًا موجزًا واضحًا، وبالتالي توضيح خصائص الظاهرة وأبعادها وما تنطوي عليه من دلالات، وهذا يمكّن الباحث من فهم العوامل الأساسية التي تؤثر في الظاهرة محل البحث، والكشف عن الفكرة الجوهرية أو القانون العام الذي يصلح لتفسير تلك الظاهرة والظواهر الأخرى ذات العلاقة بها، ومن خلال الطرق الإحصائية يتم تحديد الشروط الأساسية لموضوعية

الإجراءات المتبعة في الدراسة، كما تتيح هذه الطرق فرص اختيار بدائل متعددة من أساليب التحليل التي تناسب البيانات، ومدى إمكانية تعميم النتائج..

وإذا كان العلم في جوهره تنظيمًا موضوعيًّا للفكر، ويقوم على تبادل المعرفة بين المتخصصين في هذا المجال أو ذاك، فإن الطرق الإحصائية ذات صفة عالمية،ومن هنا تأتي ضرورة أن يكون لدى الباحثين معرفة كافية بأساليب التحليل الإحصائي وأدواته المتطورة؛ حتى يمكنهم خلق وتطوير المعرفة في مجال التخصص بما يثري الفكر الإنساني. وقد تمكنت العلوم الطبيعية- كالكيمياء والفيزياء والأحياء وغيرها- من التطور باستمرار واكتساب صفة العالمية لأسباب عديدة يأتي في مقدمتها: الارتكاز على التحليل الإحصائي والطرق الكميّة منذ عهود طويلة، وساعد على ذلك أنه في العلوم الطبيعية يمكن إخضاع الظواهر للضبط العلمي الدقيق والقياس الكمّي الصارم، وفصل المتغيرات المختلفة والتحكم في تأثير كل منها والتوصل إلى نتائج ثابتة. أما العلوم الاجتماعية والإنسانية فقد تأخرت في نشأتها الأولى عن هذا التطور لأسباب كثيرة تتعلق بطبيعة تلك العلوم؛ من حيث التعقيد واتساع مجالاتها وصعوبة إخضاع متغيراتها للتجريب والضبط العلمي الكمّي (Quantitative).

لكن الباحثين في العلوم الاجتماعية والإنسانية سرعان ما أدركوا ضرورة استخدام الإحصاء في الضبط والقياس الكمّي عند دراسة الظواهر والموضوعات، واستخدام التحليل الإحصائي والطرق الكميّة بالارتكاز على علم الإحصاء وتطبيقاته؛ الأمر الذي كان له كبير أثر في تطور العلوم الاجتماعية والإنسانية؛ من حيث طرق البحث والإجراءات، وبالتالي التوصل إلى معرفة عميقة ومنظمة.

ثانيًا:التصنيف الإحصائي

التصنيف الإحصائي هو وضع الوحدات المتجانسة أو المتشابهة – من حيث صفة معينة أو سلوك معين- في مجموعة واحدة تختلف عن غيرها من المجموعات. مثال ذلك تصنيف المبحوثين: إلى ذكور وإناث، أو إلى ريفيين وحضريين، أو إلى مشاهدين وغير مشاهدين للتليفزيون... إلخ. والتصنيف قد يكون وصفيّا فقط، بمعنى أنه لمجرد التصنيف دون أي دلالة كميّة، كما قد يكون كميّا، بمعنى أنه يوضح التصنيف ودرجة التمايز.

إن التصنيف الوصفي- كما يفهم من اسمه- هو تصنيف الوحدات دون أي دلالة كميّة للقيم المعبرة عن التصنيف، كتصنيف المبحوثين إلى ذكور وإناث من خلال تكويد الذكور بالرقم (١) والإناث بالرقم (٢)، أو تصنيفهم حسب منطقة الإقامة إلى ريف وحضر من خلال تكويد مجموعة الريف بالرقم (١) ومجموعة الحضر بالرقم (٢)؛ حيث نلاحظ أن الأرقام هنا لمجرد التصنيف فقط لأن التصنيف يهدف إلى الكشف عن وجود الصفة أو عدم وجودها.

أما التصنيف الكمّي، فهو يعكس درجة وجود الصفة أو السلوك في صورة مقدار أو كميّة (Quantity) بما يعبر عن وجود الصفة أو السلوك بدرجات معينة، من ذلك مثلًا تصنيف المبحوثين حسب مشاهدة التليفزيون إلى كثيفي المشاهدة (Heavy Viewers) ومتوسطي المشاهدة (Medium Viewers) وخفيفي المشاهدة (Light Viewers)، وذلك حسب الوقت الذي يقضيه الفرد يوميّا في مشاهدة التليفزيون من خلال تكويد خفيفي المشاهدة بالرقم (١) وتكويد متوسطي المشاهدة بالرقم (٢) وتكويد كثيفي المشاهدة بالرقم (٣). وهنا نلاحظ أن الأرقام ١، ٢، ٣ لها دلالة كميّة بجانب كونها تصنف المبحوثين من الأقل إلى الأكثر حسب معدل مشاهدة التليفزيون؛ أي أننا بصدد وحدات متمايزة من حيث الصفة والكميّة.

وسواء كان التصنيف وصفيّا أو كميّا، فإنه من أهم دعائم المعرفة البشرية؛ لأنـه يلخـص المعلومـات المتعددة في قدر صغير ذي دلالة ومعنى، كما أن التصنيف يكشـف عن الخصائص الجوهرية التي تربط الأشـياء بعضها بالبعض الآخر، ويعتمد التصنيف عـلى مـدى تمـايز الأشـياء، وعـلى تعمـيم هـذا التمايـز بحيث تنقسم الأشياء أو صفاتها إلى مجموعات بين كل مجموعة وأخرى فروق أساسية تبرر هذا الفصل القائم بينها، وتضم كل مجموعة الأفراد الذين يشتركون معًا في صفات أساسـية تـبرزهم جميعًـا معًـا في تصنيف واحد. كما يجب أن يكون أساس التصنيف منطقيّا وواضحًا، فليس من المعقـول مـثلًا تـصنيـف جمهــور التليفـزيون إلى ذكور وإناث وحضـريين، وإنما الصواب أن نقـسم الجـمهــور حـسب الجـنس (ذكور - إناث)، ثم تصنيف كل جنس حسب منطقـة الإقامة (حضـر- ريف)؛ حتى نستغـرق الأقسام الفرعية، فالذكور قد يكونون حضريين أو ريفـيين، كما أن الإناث قـد يكن حضريات أو ريفيات، وهكـذا نرى في هذا المثال أن أساس التقسيم الأول كان متغـير النوع أو الجنس، أما الأسـاس في التقـسيم الثـاني فهو متغير منطقة الإقامة.

أما حسب عدد التصنيفات، فإن التصنيف الإحصائي قـد يكـون ثنائيّـا، كـما قـد يكـون متعـددًا. إن التصنيف الثنائي يعني تصنيف الوحدات إلى قسمين فقط (كتصنيف عينة البحث إلى ذكور مقابل إناث، ريف مقابل حضر... إلخ)، أما التصنيف المتعدد فيعني تصنيف الوحدات إلى أكثر من قسمين (كتصنيف البرامج الإذاعية إلى: أخبار، منوعات، دراما، تعليمية، إعلانات، فئات)، ويمكن إحـداث تـصنيفات فرعيـة داخل كل تصنيف رئيسي أيّا كان عدد تقسيماته، من ذلك مثلًا تصنيف المبحوثين حسب الجنس (ذكـور وإناث)، ثم تصنيف كل جنس حسب السن، وتصنيف كل فئة عمرية حـسب مـشاهدة التليفزيـون..... إلخ.

ثالثًا: المتغيرات من منظور القياس الإحصائي

إن المتغير (Variable) بالمعنى البحثي والإحصائي هو الخاصية أو الصفة أو السلوك محل الدراسة، ومن منظور القياس الإحصائي تتعدد أنماط المتغيرات وفق عدة معايير أهمها:

(أ) المتغيرات النوعية مقابل المتغيرات الكميّة:

فالمتغيرات النوعية (Qualitative Variables) هي المتغيرات التي تعبر عن خاصية معينة؛ من حيث وجودها أو عدم وجودها كالجنس (ذكر، أنثى)، أو التخصص العلمي (صحافة، إذاعة وتليفزيون، علاقات عامة، تسويق، إعلان)... إلخ؛ أي أنه لا يمكن ترتيب الأفراد من الأصغر إلى الأكبر، أو من الأكبر إلى الأصغر حسب هذه الخاصية، وإنما يمكن تصنيفهم فقط.

أما المتغيرات الكميّة، فهي تلك المتغيرات التي تقاس بمقدار عددي أو كميّة (Quantity)؛ بحيث يمكن ترتيب الأفراد من الأصغر إلى الأكبر، أو من الأقل إلى الأكثر أو العكس، مثال ذلك: عدد أفراد الأسرة، عدد البرامج التليفزيونية التي يشاهدها الفرد، الوقت المنقضي في قراءة الصحيفة أو مشاهدة التليفزيون أو سماع الإذاعة، عدد أجهزة التليفزيون في المنزل، حجم الإنفاق الشهري على الإنترنت أو شراء الصحف والمجلات... إلخ.

والمتغيرات الكميّة قد تكون متصلة أو منفصلة، فالمتغيرات الكميّة المتصلة هي التي تأخذ قيمة صحيحة أو كسرية مثل: الوقت المنقضي في مشاهدة التليفزيون أو قراءة الصحف أو استخدام الإنترنت، فقد يفيد المبحوث بأنه يقضي حوالي ساعة، أو نصف ساعة أو ساعة ونصف في مشاهدة التليفزيون، فوحدة القياس هنا هي الوقت مقدّرًا بالساعة، وهو قابل للتجزئة (ساعة وربع، ساعة ونصف، ساعة وخمس دقائق... إلخ). أما المتغيرات الكمية المنفصلة فهي تلك المتغيرات التي تأخذ قيمًا صحيحة فقط مثل: عدد

أجهزة التليفزيون في المنزل، عدد الصحف اليومية التي يشترك فيها الفرد، عدد مرات التردد على السينما في العام، عدد الأفلام التليفزيونية التي شاهدها الفرد خلال الأسبوع الماضي...إلخ، فلا يمكن مثلًا أن يذكر الشخص بأن لديه جهاز تليفزيون ونصف، أو أنه يشترك في صحيفتين وربع، أو أنه شاهد ٠.٠١٢٥ فيلم تليفزيوني... إلخ، فهذه المتغيرات متغيرات كميّة منفصلة تأخذ قيمة صحيحة فقط.

(ب) المتغيرات المستقلة مقابل المتغيرات التابعة:

إن المتغير المستقل (The Independent Variable) هو المتغير الذي قد يؤثر أو يحتمل أن يؤثر في متغير آخر، أما المتغير التابع (The Dependent Variable)، فإنه المتغير الذي يحدث فيه التغير أو يقع عليه التأثير بسبب المتغير المستقل. ولتوضيح ذلك نذكر على سبيل المثال أن النوع (Gender) هو متغير قد يؤثر في اختيار التخصص الدراسي، بمعنى أن الإناث قد يكنَّ أكثر أو أقل إقبالًا على تخصصات معينة مقارنة بالذكور، وبالمثل فإن الذكور قد يكونون أكثر أو أقل إقبالًا على تخصصات معينة مقارنة بالإناث،كما أن المستوى التعليمي للأبوين قد يؤثر في الدور الوالدي، وهنا نلاحظ أن متغير النوع ومتغير المستوى التعليمي للوالدين كلاهما من المتغيرات المستقلة لأن كلّا منهما ترتب عليه تغير في متغير آخر أو يقترن بتغيرات معينة في هذا المتغير الآخر، هذا المتغير (الآخر) هو المتغير التابع (التخصص الدراسي، وكذلك الدور الوالدي).

(جـ) مستويات المتغيرات:

من منظور القياس الإحصائي،فإن المتغيرات (Variables) تصنف على النحو التالي:

١- المتغير الاسمي (Nominal):

وهو بمعنى المتغير النوعي الذي سبقت الإشارة إليه، فالمتغير الاسمي هو الخاصية أو الصفة من حيث وجودها أو عدم وجودها، كمتغير النوع (ذكور مقابل إناث)، وكذلك

متغير محل الإقامة (ريف مقابل حضر)، ومتغير الوسيلة المفضلة (الإذاعة، التليفـزيون، الصحيفة) المادة الإعلامية المفضلة (الأخبار، البرامج الإخبارية، الـدراما، الـبرامج الثقافيــة). وعنـدما نـستخدم الأرقـام في التصنيف الإحصائي لتلك المتغيرات الاسمية، فـإن الأرقـام يقتـصر دورهـا عـلى مجـرد التـصنيف دون أي دلالة كميّة، على سبيل المثال قد نرمز إلى الذكور بالرقم (١) ونرمز إلى الإناث بالرقم (٢)، فـالرقم (٢) لم يستخدم بمعنى أنه أكـبر من الرقم (١)، كما أن الرقم (١) لم يستخدم بمعنى أنه أصغر من الـرقم (٢) وإنما أُستخدم كلٌّ منهما لتصنيف الأفراد إلى ذكور وإناث.

٢- المتغير الرتبي (Ordinal):

وهو المتغير الذي يتضمن معنى المتغير الاسمي، بالإضافة إلى الدلالة الكميّة لأرقام التصنيف؛ أي أن المتغير الرتبـي يتـضمن تـصنيف الأفـراد أو الموضـوعات في مجموعـات متمايـزة، بالإضـافة إلى إمكانيـة ترتيبهم تصاعديّا أو تنازليّا، مثال ذلك متغير السن (العمر بالسنوات)؛ حيث يمكن ترتيب أفـراد البحـث حسب السن بصرف النظر عن الفروق بين هؤلاء الأفراد، فالشخص الأكبر عمرًا يأخذ ترتيبًا متقدمًا عـن ترتيب الشخص الأقل عمرًا، وهذا بدوره يأخذ ترتيبًا متقدمًا عن الشخص الأصغر منـه... وهكـذا. لكـن ذلك لا يعني تساوي الفروق في الأعمار بين المفحوصين، فقد يكون الشخص الأكـبر سـنّا عمـره ٣٧ سـنة، والذي يليه في الترتيب عمره ٣٤، أما الذي يليه فقد يكون عمره ٢٦ سنة، والـذي يليـه ١٥ سـنة... إلـخ، فالفروق في الأعمار غير متساوية.

وفي البحوث الإعلامية تستخدم المتغيرات الرتبية بكثرة، مثل ترتيب المبحـوثين حـسب عدد الساعات التي يقضونها يوميا في مشاهدة التليفزيون أو استخدام الإنترنـت أو قـراءة الصحف... إلـخ؛ بحيـث يكـون لكـل مبحـوث رتبـة معينـة حـسب الوقـت المنقـضـي في مشاهدة التليفزيون أو استخدام الإنترنت، كـما يمكـن تـصنيف المبحـوثين وتـرتيبهم مـن

الأصغر للأكبر (أو من الأكبر للأصغر)، بصـرف النظر عن الفروق بينهم في الوقت المستغرق في مشاهدة التليفزيون أو في استخدام الإنترنت، فقد تكون هذه الفروق كبيرة أو صغيرة.

٣- المتغير الفئوي أو الفتري (Interval):

وهو يعبر عن كميّة أو مقدار؛ أي أن له وحدة قياس تعكسها الأعداد من الصفر إلى مالا نهاية، كـما أن المتغير الفئوي (الفتري) يتضمن معنى المتغير الرتبي، بالإضافة إلى تساوي الفروق بين الوحدات التـي يتم تصنيفها، كأن يتم تصنيف المبحوثين حسب السن إلى:

- من ٢٠ إلى ٢٩ سنة
- من ٣٠ إلى ٣٩ سنة
- من ٤٠ إلى ٤٩ سنة

فالفرق بين كل مجموعة هو عشر سـنوات؛ أي أن المبحـوثين تـم تـرتيبهم في مجموعـات، والأخـرى الفروق متساوية بين هذه المجموعات.

وتستخدم المتغيرات الفترية كثيرًا في البحوث الإعلامية، من ذلك مثلًا متغير الوقت اليومي المنقـضي في مشاهدة التليفزيون أو الوقت اليومي المنقضي في سماع الراديو أو قراءة الـصحف، درجـة المبحـوثين على مقياس الاستخدامات والإشباعات في علاقتهم بالوسيلة، أعمار المفحوصـين، قيمـة الـدخل، الدرجـة على مقياس المستوى الاقتصادي الاجتماعـي، عـدد مـرات الـتردد عـلى الـسينما أو المـسـرح خـلال فـترة معينة، عدد الأفلام السينمائية، أو عدد البرامج التـي يتعـرض لهـا المبحـوث، عـدد الأخبـار المتـضمنة في النشـرة، مدة البرامج، مساحة المواد الصحفية بالسنتيمتر المربع... إلخ. كل هذه أمثلـة لمتغـيرات فتريـة أو فاصـلة تتردد كثيرًا في بحوث الإعلام.

وإذا كان المتغير الفئوي أو الفتري له وحدة قياس تبدأ من الصفر إلى مالا نهاية، فإن البداية الصفرية تنطبق على العلوم الطبيعية (كالكيمياء والفيزياء والرياضيات)؛ إذ إن (الصفر) يعني الغياب الكامل للظاهرة أو الموضوع (صفر مطلق). أما في العلوم الإنسانية بما في ذلك علم الاتصال، فإن (الصفر) لا يعني الغياب الكامل للظاهرة أو الخاصية المقاسة، أو السلوك، فعندما يفيد بعض المفحوصين بأنهم «لا يشاهدون التليفزيون»، فإن هذا لا يعني أنهم لم يشاهدوا التليفزيون طوال حياتهم، وإنما هم شاهدوه بالتأكيد، وعندما يحصل بعض الطلاب على (صفر) في مقرر الإحصاء، فإن هذا لا يعني أنهم لا يعرفون شيئًا على الإطلاق في هذا المقرر، وإنما هم بالتأكيد يعرفون بعض الأشياء، فالصفر هنا (صفر نسبي) أي غير مطلق.

عند هذا المستوى من المتغيرات - وأقصد بها المتغيرات الفئوية أو الفاصلة، تعمل البحوث في مجالات العلوم الاجتماعية والإنسانية عمومًا، فهذه العلوم لم تعتمد بعد على المتغير النسبي (Ratio) على النحو المعمول به في العلوم الطبيعية.

٤- المتغير النسبي (Ratio):

وهو المتغير الذي تتوفر فيه معاني المتغير الفتري بجانب كونه يتضمن الصفر المطلق، بمعنى غياب الصفة المقاسة، فعندما نقول إن درجة الحرارة تساوي (الصفر)، فهذا يعني الغياب الكامل للحرارة، وبوجه عام، فإن المتغيرات النسبية (Ratio Variables) تقع ضمن العلوم الطبيعية، وليس العلوم الاجتماعية بما في ذلك علم الاتصال.

ومن الضروري أن يكون الباحث على دراية كافية بخصائص المتغيرات؛ لأنه في ضوء هذه الخصائص يتم وضع نظام إدخال البيانات واختيار المعالجة الإحصائية المناسبة، كمثال على ذلك، فإن اختيار معامل الارتباط المناسب يتطلب تحديد نوعية المتغيرات المطلوب حساب الارتباط بينها؛ أي ما إذا كانت تلك المتغيرات اسمية (Nominal) أو

رتبية (Ordinal) أو فئوية (Interval) (حيث يشيع استخدام هذه الأنماط الثلاثة من المتغيرات في العلوم الإنسانية).

وعندما نكون بصدد تقصي الارتباط بين متغيرين (Two variables)، فإن هذا يعني أن الارتباط قد يكون بين متغيرين اسميين، أو بين متغيرين رتبيين، وقد يكون بين متغيرين فتريين، كما قد يكون الارتباط بين متغير اسمي والآخر رتبي، أو بين متغير اسمي والآخر فتري، وقد يكون الارتباط بين متغير فئوي (فتري) والآخر رتبي، ولكل نمط من هذه العلاقات معامل الارتباط الذي يناسبها، بمعنى معامل الارتباط المناسب لطبيعة المتغيرين المطلوب رصد العلاقة بينهما (فلا يصح مثلًا استخدام معامل ارتباط بيرسون بين متغيرين اسميين)، وذلك على النحو الذي يوضحه الفصل الخاص بمعاملات الارتباط.

(د) أنماط أخرى من المتغيرات:

وإذا كانت المتغيرات من منظور القياس الإحصائي تصنف إلى اسمية ورتبية وفترية ونسبية، فإن هناك أنواعًا أخرى من المتغيرات التي يجب أن يعيها الباحث، ومن أهمها: المتغير الدخيل، المتغير الرمزي، المتغير الوسيط.

(١) المتغير الدخيل (Extraneous Variable)، ويسمى أحيانًا المتغير المحير أو المربك (Confound Variable) وهو المتغير الذي لا يتم التحكم فيه ضمن الدراسات التجريبية، ويؤثر في المتغير التابع رغم أن التأثير يفترض أن يكون بفعل المتغير المستقل. ففي الدراسات التجريبية يتعين التحكم في جميع المتغيرات محتملة التأثير في المتغير التابع (Dependent Variable) حتى نضمن أن يكون التغير في المتغير التابع ناتجًا عن المتغير المستقل (Independent Variable) فقط، غير أنه في بعض الأحيان يحدث تغيير في المتغير التابع ليس بسبب المتغير المستقل؛ ولكن بسبب وجود متغير آخر لم يتم التحكم فيه، هذا المتغير (الآخر)، هو ما يعرف بالمتغير الدخيل، إنه يرتبط بالمتغير المستقل، وعندما يتأثر

المتغير التابع، فإن هذا التأثير يكون محل تساؤل: هل هو بسبب المتغير المستقل فقط كما هو مفترض؟ أم أن هذا التأثير بسبب المتغير الدخيل فقط؟ أم أن هذا التأثير بسبب التفاعل بين المتغير المستقل والمتغير الدخيل؟ إن البحث قد يجد علاقة جوهرية بين استخدام وسائل الإعلام (متغير مستقل) والمعرفة السياسية (متغير تابع)، في الوقت الذي تكون هذه العلاقة بسبب متغير آخر يرتبط بمتغير استخدام وسائل الإعلام، هذا المتغير الآخر قد يكون مستوى التعليم، أو الاهتمام بالسياسة، أو العضوية في حزب سياسي..... إلخ.

(٢) المتغير الرمزي (Dummy Variable)، وهو المتغير المطلوب التنبؤ به أو توقع حدوثه (أو عدم حدوثه) في تحليل الانحدار- خاصة الانحدار اللوجستي- ويتم تكويد هذا المتغير بالرقم (صفر) والرقم (١)، ويدل الرقم (صفر) على عدم حدوث المتغير، بينما يدل الرقم (١) على حدوث المتغير، فإذا افترضنا أن المتغير الرمزي هو التصويت في الانتخابات، فإن المشاركة في التصويت تأخذ الرقم (١). أما عدم المشاركة في التصويت فيأخذ الرقم (صفر)، هذا المتغير هو المتغير الرمزي (Dummy Variable). ويكون هناك مجموعة من المتغيرات المستقلة يتم تقصي أثرها في ترجيح حدوث التصويت في الانتخابات.

(٣) المتغير الوسيط (Mediating Variable)، وهو المتغير الذي يؤثر في العلاقة بين المتغير المستقل والمتغير التابع، فقد يجد الباحث أن هناك علاقة بين السن واستخدام مواقع التواصل الاجتماعي، لكن هذه العلاقة تتوقف على مستوى التعليم، وهنا يكون مستوى التعليم هو المتغير الوسيط، وهو من ضمن المتغيرات التي تتناولها الدراسة.

رابعًا: الإحصاء الوصفي والإحصاء الاستدلالي

الإحصاء الوصفي (Descriptive Statistics)، هو كل المقاييس الإحصائية التي تعتمد على البيانات المستمدة من العينة (Sample) في وصف الظاهرة أو السلوك أو الموضوع أيًا كانت طبيعة هذا الوصف، فالباحث الذي يدرس علاقة عينة من المفحوصين بالفضائيات

مثلًا يقوم بجمع البيانات وتحليلها وتنظيمها وجدولتها؛ بحيث تتضمن المعطيات الإحصائية التي «تصف» هذه العلاقة، قد تكون هذه المعطيات في صورة تكرارات ونسب مئوية، أو متوسطات وانحرافات معيارية ومقارنات... إلخ؛ أي أن المعطيات الإحصائية هنا تظهر خصائص علاقة عينة الدراسة بالقنوات الفضائية مجال البحث.

أما الإحصاء الاستدلالي (Inferential Statistics)، فيعني التعميم من الجزء على الكل، بمعنى استخدام النتائج المستمدة من دراسة العينة في التعميم على المجتمع، مثال ذلك أن يقوم الباحث بدراسة على عينة عشوائية من طلاب الجامعات؛ للتعرف على نسبة قراءة الصحف، وبناءً على هذه النسبة (المستمدة من العينة)، يقوم بتقدير نسبة قراءة الصحف في المجتمع ككل (طلاب الجامعات)، فإذا كان ٢٩% من العينة يقرؤون الصحف اليومية، فإن هذه النسبة يمكن من خلالها تقدير نسبة قراءة الصحف اليومية في المجتمع، وإذا كان متوسط درجة العينة على مقياس الاستخدامات والإشباعات فيما يخص قراءة الصحف اليومية هو ١٦ بانحراف معياري ٢.٥، فإنه يمكن الاستفادة من هذه المعطيات في تقدير متوسط المجتمع..... بوجه عام، فإن الإحصاء الاستدلالي هو تقدير الظاهرة في المجتمع بناء على نتائج العينة (انظر المبحث الثالث في الفصل السادس).

وسواء كان الإحصاء وصفيًا أو استدلاليًا، فإن استخدامه في تحليل البيانات يتم في إطار المتغيرات المطلوب قياسها، والمقارنات المطلوب إجراؤها؛ وذلك بهدف تيسير إمكانية التوصل إلى تفسيرات صادقة ومفيدة ومختصرة من بيانات كثيرة، ولكي يستخدم الإحصاء الوصفي أو الإحصاء الاستدلالي استخدامًا مناسبًا في تحليل البيانات، يجب أن ترتكز عملية جمع البيانات ومعالجتها وتنظيمها وتفسيرها على التساؤلات التي يريد الباحث الإجابة عنها، أو الفروض التي يريد التحقق منها، مع الأخذ في الاعتبار عشوائية العينة وطبيعة البيانات المستخدمة ومستويات المتغيرات المدروسة؛ إذ إن عدم مراعاة هذه الاعتبارات الأساسية هو الذي يؤدي إلى إساءة استخدام الإحصاء في البحوث والتوصل إلى نتائج زائفة وتفسيرات مضللة.

خامسًا: الإحصاء البارامتري والإحصاء اللابارامتري

الإحصاء البارامتري أو الإحصاء المعلمي (Parametric Statistics)، يشمل كافة الطرق الإحصائية التي تستند على معرفة خصائص المجتمع الذي سحبت منه العينة، من هذه الخصائص مثلًا أن يكون توزيع الظاهرة في المجتمع الذي سحبت منه العينة توزيعًا اعتداليًا (Normal Distribution)، ومن الثابت أن الطرق البارامترية أو المعلمية تشترط الاختيار العشوائي للعينة، كما تشترط توفر معلومات عن توزيع المجتمع حسب المتغيرات التي نهتم بدراستها، وتستخدم طرق الإحصاء البارامتري مع العينات الكبيرة، كما أنها تناسب البيانات الفئوية والبيانات النسبية (Ratio). ومن أمثلة الأساليب الإحصائية البارامترية نذكر: معامل الارتباط الثنائي، معامل الارتباط الثنائي الأصيل، معامل ارتباط بيرسون، الأساليب الإحصائية التي ترتكز على استخدام الوسط الحسابي من الدرجات الخام (كما هو الحال في اختبار T وتحليل التباين..).

أما الإحصاء اللابارامتري (Non-Parametric Statistics)، فهو لا يشترط خصائص معينة في المجتمع الذي سحبت منه العينة، ولا يشترط اعتدالية توزيع الظاهرة في المجتمع، كما لا يشترط الاختيار العشوائي للعينة. وتصلح الطرق اللابارامترية غالبًا للعينات الصغيرة (وإن كانت في بعض الأحيان تستخدم مع العينات الكبيرة)، كما أنها تناسب البيانات الاسمية والرتبية (وفي بعض الأحيان تستخدم مع البيانات الفئوية والبيانات النسبية). ومن أمثلة الأساليب الإحصائية اللابارامترية تلك الأساليب التي تستخدم للتحقق من الفروض الارتباطية (التي تختبر العلاقة أو الارتباط)، ويدخل في عداد ذلك: معامل ارتباط سبيرمان، معامل ارتباط كندال، معامل التوافق، معامل فاي، معامل جاما، معامل لامدا... إلخ)، وهناك أيضًا أساليب إحصائية لابارامترية للتحقق من الفروض الفارقة، وهي الفروض التي تختبر الفروق، وذلك اعتمادًا على الرتب (Ranks) (وليس على الدرجات الخام)، ويدخل في عداد ذلك: اختبار كولموجروف-سميرنوف، اختبار مان- ويتني، اختبار كروسكال واليس، اختبار فيشر، اختبار الوسيط، اختبار ماكنمار.... إلخ.

المبحث الثاني

الرموز والعمليات الجبرية الأساسية في الإحصاء

أولاً: الرموز الإحصائية

نستخدم في الرياضيات والإحصاء الكثير من العلامات والرموز والاختصارات والإشارات، ومعظمها يستخدم عالميًا، ومن أمثلة ذلك: علامة الجمع (+)، وعلامة الطرح (-)، وعلامة القسمة (÷)، وعلامة الضرب (×).

وفي عمليات الضرب يمكن استخدام الأقواس بدلًا من علامة الضرب، فمثلًا:

٢ × ٩ يمكن كتابتها (٢) (٩)، أو ٢(٩) وبذلك تحل الأقواس محل علامة الضرب.

وفي حالة استخدام الحروف مثل س، ص مثلًا للدلالة على متغيرات معينة، يمكن التعبير عن علامة الضرب بنقطة (س . ص)، كما يمكن كتابة الحرفين فقط (س ص)، وفي كلتا الحالتين فإن المعنى يكون (س × ص).

وفي عمليات القسمة يمكن التعبير عن إشارة القسمة بصور مختلفة، فمثلًا:

٥ ÷ ١٠ يمكن كتابتها ٥ / ١٠ أو $\frac{5}{10}$

وبذلك تتغير صورة إشارة القسمة ولكن معناها ونتائجها لم تتغير.

وهناك إشارة = (يساوي)، فمثلًا ١٢ – ٤ = ٨، وكثيرًا ما تظهر إشارة = (يساوي) بصورة مختلفة فنجدها ≅، وهذه الإشارة ≅ تدل على أن النتائج تقريبية، فمثلًا ٥.٥ × ٥.٥ ≅ ٣٠.٣؛ أي الناتج تقريبًا يساوي ٣٠.٣ (إذ إن الناتج الحقيقي هو ٣٠.٢٥).

وهناك إشارة ≠ وهي تعني (لا يساوي)؛ أي أنها تعني أن القيمتين غير متساويتين، فمثلًا ١٠ + ٥ ÷ ٢ ≠ ١٠ + ٢ ÷ ٥ وكذلك:

س ≠ ص؛ أي أن س لا يساوي ص

وهناك أيضًا إشارة ≡، ومعناها (تطابق) بمعنى أن المقدارين متطابقان، فمثلًا (س + ص)^{٢} ≡ س^{٢} + ٢س ص + ص^{٢} ويسمى هذا الناتج «مفكوك المربع الكامل» لأن (س+ ص)^{٢} = س^{٢}+ س ص+ ص س + ص^{٢}

ومن الشائع استخدام الاختصارات في الإحصاء؛ توفيرًا للوقت وتحقيقًا للدقة في الصورة الرياضية، أو الإحصائية.. ومن هذه الاختصارات:

> وتعني (أكبر من)

< وتعني (أصغر من)، أو (أقل من)

≤ وتعني (أكبر أو تساوى)

≥ وتعني (أصغر من أو يساوى)

فمثلًا: س < ص (بمعنى أن س أقل من ص)، وكذلك س > ٩ بمعنى أن س أكبر من ٩، وبالمثل س < ٩ تعني أن القيمة (س) أصغر من ٩

س ≥ ٩ تعني أن س أقل من أو يساوي ٩

وكذلك س ≤ ٩ معناها أن القيمة (س) أكبر من أو يساوي ٩

ومن الاختصارات الإحصائية أيضًا df وهو اختصار لمصطلح درجة الحرية (Degree of Freedom) وفي اللغة العربية فإن اختصار درجة الحرية هو (د ح)، فهناك حد إحصائي في مقام تباين العينة وهو (ن – ١)، أو (n-1). وتعني الحدود المسموح بها أو

المتاحة، كمثال توضيحي نفرض أن المطلوب اختيار ثلاثة أعداد. في هذه الحالة يمكن اختيار أي ثلاثة أعداد (ولتكن ٣، ٥، ٢)؛ أي أنه لا توجد أي قيود أو اشتراطات أو مواصفات للأعداد التي نختارها. لكن إذا كان المطلوب اختيار ثلاثة أعداد بحيث يكون متوسطها ٥، فإن هناك شرطًا أو قيدًا في الاختيار، إننا يمكننا اختيار العدد ٣ والعدد ٥، أما العدد الثالث فلابد أن يكون مكملًا للمجموع بحيث يكون متوسط هذا المجموع هو ٥؛ أي أن لدينا الحرية في اختيار عددين فقط، أما العدد الثالث فلابد من اختياره بمواصفات معينة. أي أنه إذا كان المطلوب اختيار عدد n من الأعداد بحيث يكون لها متوسط محدد، فإن درجة الحرية تكون (n-1) وأن آخر عدد نختاره يجب أن يكون مناسبًا بحيث يمكن الحصول على المتوسط المطلوب.

وتستخدم الرموز والحروف اللاتينية كثيرًا في الإحصاء لوصف الطرق والمعطيات الإحصائية، وعادة ما تدل الرموز الفنية على المؤشرات والإحصاءات، ومن المتعارف عليه عالميّا أن المؤشرات (أو المعالم) الخاصة بالمجتمع تمثل بحروف يونانية، أما الإحصاءات فيرمز لها بالحروف الرومانية، من ذلك مثلًا:

يستخدم الحرف μ بمعنى متوسط المجتمع، وهو حرف لاتيني ينطق (ميو)، أما متوسط العينة فيرمز له بالحرف $\overline{X}$ (حرف X فوقه شرطة معتدلة)، وتنطق (إكس بار).

كما أن الحرف N يرمز إلى عدد المجتمع، أما العينة فيرمز لها بالحرف نفسه ولكن بالصورة الصغيرة n.

ويستخدم الحرف اللاتيني π (باي) بمعنى النسبة في المجتمع، بينما يستخدم الحرف P ليرمز إلى نسبة المفردات التي لها صفة معينة في العينة، كما يستخدم الحرف نفسه p ليشير إلى الاحتمال، كاختصار لكلمة (Probability).

ومن الشائع أيضًا استخدام الحرف اللاتيني ألفا α ليدل على المعنوية (Significance)، وإن كانت له استخدامات أخرى- حسب السياق الوارد فيه.

ويرمز لتباين المجتمع بالرمز σ^2 (سيجما تربيع)، أما تباين العينة فيرمز له بالحرف S^2؛ ونظرًا لأن التباين هو مربع الانحراف المعياري، فإن الانحراف المعياري للمجتمع يرمز له بالحرف σ سيجما، أما الانحراف المعياري للعينة فيرمز له بالحرف S.

وهناك الحرف $\sum$ ويعني مجموع، وينطق سيجما، وأحياناً يستخدم الحرف $\sum$ كصورة مكبرة للحرف σ.

وهناك الرمز ! وينطق فكتوريال، بمعنى أن ! س (تنطق فكتوريال س)، كما أن ! ٦ = ٦ × ٥ × ٤ × ٣ × ٢ × ١ = ٧٢٠

ومن الرموز المعروفة التي تستخدم في الإحصاء والرياضيات أيضًا، إشارة: ∞ أو - ∞ ومعناها إلى مالا نهاية موجب أو سالب.

وإذا كان ذلك فيما يخص الأعداد (Numbers)، فإن هناك إشارات تستخدم مع المجموعات (Groups)، أهم هذه الإشارات $\cap$ ومعناها تقاطع مفردات كل من س، ص، وعكسها إشارة $\cup$، ومعناها اتحاد أو تجانس مفردات كل من س، ص، وكذلك إشارة ناقص (-) ومعناها الفرق. فمثلًا: {٢، ٣} تسمى مجموعة، وكذلك {٤، ٥، ٦} تسمى مجموعة.

وفيما يخص التقاطع $\cap$، فإنه يعني العدد أو الأعداد الموجودة في المجموعة الأولى، وفي الوقت نفسه موجودة في المجموعة الثانية، فمثلًا:

{٢، ٣} $\cap$ {٣، ٥} = {٣} بمعنى أن الرقم ٣ موجود في المجموعة الأولى والمجموعة الثانية، أي أن هاتين المجموعتين تتقاطعان في الرقم ٣

أما فيما يخص الاتحاد $\cup$ فإن معناه جميع العناصر الموجودة في المجموعة الأولى وجميع العناصر الموجودة في المجموعة الثانية، فمثلًا:

{٢، ٣} ∪ {٣، ٥} = {٢، ٣، ٥} بمعنى أن الأرقام ٢، ٣، ٥ تشترك فيها المجموعة الأولى والمجموعة الثانية.

أما فيما يخص الفرق، ورمزه علامة ناقص (-)، فإنه يعني أن العدد موجود في المجموعة الأولى وغير موجود في المجموعة الثانية.

{٢، ٣} - {٣، ٥ } = {٢} بمعنى أن الرقم ٢ موجود في المجموعة الأولى وغير موجود في المجموعة الثانية.

ثانيًا: إطلالة على أهم العمليات الجبرية للإحصاء

إن استخدام الإحصاء في تحليل البيانات يحتاج إلى قدر من التفكير المنطقي في المشكلة البحثية وكيفية معالجتها إحصائيًا، ويمكن للباحث أن يتقن ذلك سواء كان لديه خلفية قوية في الرياضيات أم لا، فالمهم أن يكون لديه الرغبة في فهم الأساليب الإحصائية التي يمكن أن تساعده في تحليل بيانات بحثه للتوصل إلى نتائج سليمة، كما أن عملية تحليل البيانات تتطلب قدرًا من العمليات الحسابية والجبرية، فالإحصاء يرتكز على أسس رياضية يتسع مداها حتى ليكاد يشمل الرياضيات بفروعها المختلفة بما يخرج عن نطاق اهتمام الجزئية التي نناقشها. لكن هناك بعض العمليات الجبرية الأساسية التي تفيد الدارسين والباحثين رغم بساطتها:

فعندما تحتوي مسألة من المسائل على عمليات حسابية مختلفة (الجمع، الطرح، والضرب، والقسمة)، ينبغي إجراء عمليتي الضرب والقسمة أولًا، وذلك حسب ترتيب كل منها، بمعنى أنه إذا أتت إشارة الضرب أولًا نقوم بعملية الضرب ثم القسمة، أما إذا أتت إشارة القسمة أولًا، فإننا نقوم بعملية القسمة أولًا ثم الضرب. فمثلًا:

$$٢٠ - ٦ \times \frac{9}{3} = ٢$$

ذلك أننا بدأنا بضرب الرقم (٦) في ناتج قسمة الرقم ٩ على ٣ وبما أن ٩÷ ٣ = ٣ فيكون: ٦ × ٣ = ١٨ وبالتالي فإن ٢٠ – ١٨ = ٢ أي أن عملية القسمة وعملية الضرب تمت قبل عملية الطرح.

وكذلك: ٧ + ٦ × ٤ - ٢ × ٣ = ٢٥

لقد بدأنا بضرب ٦×٤ وهي تساوي ٢٤

ثم ضرب ٢ × ٣وهي تساوي ٦

فأصبح لدينا ٧ +٢٤- ٦ فيكون الناتج ٢٥

أي أن عملية الضرب تمت قبل كل من عملية الجمع وعملية الطرح. الاستثناء من هذه القاعدة هو عندما تأتي الأرقام داخل الأقواس:

فمثلًا (٨+٥) × (٦+١) = ١٣ × ٧ = ٩١

وكذلك:

(٣ + ٢) × $\frac{2}{3}$ ÷ (٦ + ١)

= ٥×٠.٦٧ ÷ ٧

= ٣.٣٥ ÷ ٧ = ٠.٤٧٨٥٧

لقد تم التعامل مع الأرقام الموجودة داخل الأقواس على أنها مستقلة؛ أي بمفردها، ويجب فك هذه الأقواس أولًا.

وعندما يتم فك الأقواس وتكون الإشارة السابقة للقوس سالبة، فإن كل الإشارات داخل هذا القوس تتغير إلى العكس، بمعنى أن إشارة + تصبح – وإشارة – تصبح + فمثلًا:

(٩ + ١ – ٤) - (٦ + ٢ – ٣) = ٩ + ١ – ٤ – ٦ -٢ + ٣ = ١

فالإشـارات الخاصة بالأرقام في القـوس الثاني تغـيرت إلى العكـس، فقـد كانـت (٦+٢-٣)؛ ونظـرًا لأن القوس تسبقه إشارة -، فإن الأرقام المذكورة أصبحت (-٦-٢+٣)

وبالإضافة إلى إجراء عمليات القسمة والضرب أولًا إذا كانت المسألة الحسابية المراد حلهـا تتـضمن إشارات الجمع والطرح والقسمة والضرب، فإن القواعد الجبرية تقضـي بالتعامل مع الأعداد التي تسبق هذه العمليات على أنها مستقلة كما سبق القول، فمثلًا:

٢ + ٢٥ ÷ ٥ × ٣ = ١٧

في هذا المثال بدأنا بقسمة ٢٥÷٥ فيكون النـاتج ٥، ثـم تـم ضرب ٥×٣ فيكـون النـاتج ١٥، وأخـيرًا أضفنا ١٥ إلى ٢. لقد تم التعامـل مـع العـدد ٢ كعـدد كامـل ومـستقل لأنـه يـسبق الأعـداد المطلـوب قسمتها والأعداد المطلوب ضربها، أما الأعداد (٢٥ ÷ ٥ × ٣) فهي مجموعة واحدة، ويجب الانتهـاء أولًا من عمليتي الضرب والقسمة لهذه الأعداد كـما قلنـا سـابقًا؛ أي أن (٢٥ ÷ ٥ × ٣) = ١٥، ثـم نـضيف الناتج على + ٢ فتصبح ٢ + ١٥ = ١٧

ومن البديهيات الرياضية للإحصاء ما يتعلق بالتعامل الجبري مع الأعداد الـسالبة والموجبـة، مـن ذلك مثلًا أنه عند جمع الأعداد ذات الإشارات الموحدة، تكون عملية الجمع متفقة مـع تلـك الإشـارات، وبالتالي تكون عملية سهلة، أما في عملية جمع الأعداد ذات الإشارات المختلفة، فإنه يتم جمـع الأعـداد ذات الإشارات الموجبة، ثم جمع الأعداد ذات الإشارات السالبة، ويحمل الناتج إشارة أكبر قيمة، فمثلًا:

(+٨) + (+٦) + (-٣) + (-٦) = ١٤ - ٩ = + ٥

وكذلك:

+٨ (-٤) + (-٥) = ٨- ٤- ٥= -١

كما أن:

(+٢) + (-١٠) + (-٩) + (+٧) = - ١٩ + ٩ = - ١٠

ومن المفيد هنا معرفة أن ناتج ضرب الأعداد متشابهة الإشارات يكون الناتج موجبًا، أما عند ضرب الأعداد مختلفة الإشارات فيكون الناتج سالبًا، فمثلًا:

+٥ × +٦ = +٣٠

-٥ × -٦ = +٣٠

+٥ × -٦ = -٣٠

-٥ × +٦ = -٣٠

وفي حالة عدم وجود إشارة الجمع أو الطرح، فإن الرقم يكون موجبًا، فمثلًا:

(٥) × (٦) هي نفسها (+٥) × (+٦)

٥ × -٦ هي نفسها (+٥) × (-٦)

في هذا السياق أيضًا فإن جمع أو طرح الصفر إلى أو من أي عدد، أو رمز، لا يغير العدد، أو الرمز، فمثلًا ٥+ صفر= ٥، ٥- صفر= ٥، س + صفر= س، س- صفر= س

أما إذا ضرب الصفر في أي عدد، أو رمز، فإن الناتج يكون صفرًا، وكذلك الحال إذا قسمنا الصفر على أي عدد، أو رمز، فإن النتيجة تساوي صفرًا أيضًا (وإن كان من المعتاد أن الصفر لا يستخدم مقسومًا عليه)، وعلى الرغم من أن وجود الصفر على يسار أي رقم صحيح لا يكون له قيمة، إلا أنه يمكن استخدام الصفر على يسار الأرقام الصحيحة وتكون له قيمة أساسية، فإذا كان على يسار الرقم ١ ستة أصفار مثلًا، فإن ذلك قد يعني الواحد بعد المليون، وإذا كان على يسار الرقم ١ ثلاثة أصفار، فإن ذلك قد يعني الواحد بعد الألف..... وهكذا.

المسألة الأخرى ضمن البديهيات الجبرية للإحصاء هي الجذر التربيعي؛ إذ إن الجذر التربيعـي لأي رقم هو ذلك العدد الذي إذا ضُرب في نفسه نتج هذا الرقم. وكثيرًا ما تستخدم في الإحصاء إشارة الجذر $\sqrt{\ }$ وهذه الإشارة إذ لم تميز بأي رقم يشير إلى قوتها فإنها تعبر عن الجذر التربيعي. فمثلًا:

$$\sqrt{81} = 9$$

وفي هذه الحالة يسمى العدد ٨١ المجذور (Radicand). أما العدد ٩ فيسمى الجذر (Root).

ولا يوجد جذر تربيعـي حقيقي للقيم السالبة؛ أي التي تسبقها علامة ناقص (-)، فالعـدد -٨١ مـثلًا ليس له جذر تربيعي حقيقي.

وعندما نقوم بضرب الجذور التربيعية، فإن الجذر التربيعي لكل عدد على حدة مـضروبًا في الجـذر التربيعي للعدد الآخر يكون مساويًا للعددين مضـروبين في بعضهما تحت الجذر، فمثلًا:

$$\sqrt{4} \times \sqrt{9} = \sqrt{9\times4}$$

أما في حالة القسمة فنجد أن $\sqrt{36} \div \sqrt{4} = \sqrt{\frac{36}{4}}$

وفي حالة الضرب أيضًا، فإن أحد العددين قد يكون غير مجـذور بيـنما العـدد الثـاني مجـذور، وفي هذه الحالة يتم الحصول على نتاتج الضـرب بتربيع العدد الأول ووضع العددين تحت الجذر، فمثلًا:

$$9 \times \sqrt{36} = \sqrt{36\times81}$$

وفي حالة القسمة فإن:

$$٣ = \overline{\frac{27}{3}}\ b = \frac{9}{3}$$

وقبل البدء في عمليات الجمع أو الطرح يجب الحصول أولًا على الجذور، فمثلًا:

$$٥ = \overline{9}\ b + \overline{4}\ b$$

وقيمة هذا الناتج لا تساوي ناتج $\overline{9+4}\ b$

هناك أيضًا الأسس واللوغاريتمات باعتبارها ضمن البديهيات الجبرية للإحصاء. فمن المعروف أن «أس» (Exponent) أي رقم هو عدد المرات التي يكون فيها هذا الرقم مضروبًا في نفسه للحصول على قيمة معينة، ويشترط في الأسْ أن يكون عددًا صحيحًا موجبًا. فالمقدار ص٥ (ص أس خمسة) يدل على أن ص مضروب في نفسه خمس مرات، وهكذا فإن س٤ = س × س × س × س، كما أن $٧^{٣}$ (وتنطق ٧ أس ٣)، تساوي ٧ × ٧ × ٧ = ٣٤٣

وكذلك $٢^{٣} \times ٢^{٥}$ = (٢×٢×٢) × (٢×٢×٢×٢×٢) = $٢^{٣+٥}$ = $٢^{٨}$ (٢ أس ثمانية)

كما أن $٧^{٢} \times ٧^{٣} \times ٧^{٤}$ = $٧^{٢+٣+٤}$ = $٧^{٩}$ (سبعة أُسْ تسعة).

أما اللوغاريتمات (Logarithms)، فهي مجال متسع وتخصص قائم بذاته، ويستخدم الاختصار «لوج» (Log) كمختصر لكلمة (لوغاريتم)، بمعنى أن لوج س (معناها لوغاريتم س)، ويستخدم الحرف الإنجليزي e بمعنى اللوغاريتم الطبيعي وقيمته ٢.٧١٨٢٨

وبوجه عام، فإن لوغاريتم أي عدد معلوم هو «الأس» الذي يرفع إليه الأساس لينتج العدد المفروض، ويسمى العدد الأصلي الأساس، أما الرقم الذي يرفع إليه هذا العدد

فيسمى «الأس». على سبيل المثال فإن $٥^{٣}$ = ١٢٥؛ أي أن العدد ٥ يجب أن يرفع إلى الرقم ٣ حتى نحصل على ١٢٥ وقد اتخذ الرقم ١٠ أساسًا عامّا لتسهيل عمليات الحساب من جداول اللوغاريتمات. وفي مجال الإحصاء تستخدم جداول اللوغاريتمات التي تتخذ الأساس الخاص بها العدد ١٠. فمثلًا، ١٠٠٠ = $١٠^{٣}$، وتكتب لوج ١٠٠٠ = ٣، كما أن ١٠٠ = $١٠^{٢}$؛ أي لوج ١٠٠ = ٢، وكذلك ١٠ = $١٠^{١}$، وتكتب لوج ١٠ = ١.. وهكذا.

المبحث الثالث

استخدام الحاســـب الآلي في التحليل الإحصائي

المطلب الأول: الحاسب والبرامج الإحصائية

أولًا: مزايا الحاسوب وحدوده

أدى التقدم الكبير الذي حدث في الآلات الحاسبة والحواسيب إلى تسهيل وسرعة إجراء العمليات الحسابية مهما كانت معقدة.المشكلة أن الكثير من الباحثين لا يعطون الاهتمام الكافي لتحليل البيانات الميدانية بحجة أن الحاسوب (سيجري) هذا التحليل عن طريق البرامج الإحصائية المتوافرة! لكن من الضروري هنا التأكيد على أن عملية تحليل البيانات تعد جزءًا أساسيّا لا يتجزأ من عملية البحث ذاتها، ويترتب عليها استخلاص النتائج والتفسيرات.

ومما لا شك فيه أن الحاسوب يمكنه تخزين كميات كبيرة من البيانات يعجز العقل البشري عن تخزينها والتعامل معها، كما يمكن للحاسوب إجراء العديد من العمليات الإحصائية والرياضية المعقدة على البيانات التي يحتاج العقل البشري إلى وقت طويل

وجهد شاق لإنجازها، فالحاسوب يتميز بالسرعة الفائقة والدقة المتناهية في إجراء هذه العمليات؛ مما يوفر على الباحث كثيرًا من الوقت والجهد.

ومن المعروف أيضًا أن الباحث في العلوم الاجتماعية بعامة لا يمكنه دراسة الظواهر السلوكية دراسة متعمقة دون أن يأخذ في الاعتبار المتغيرات الأساسية التي يمكن أن تؤثر في هذه الظواهر، وهذا يجعل البيانات التي يجمعها الباحث عن ظاهرة معينة متعددة المتغيرات؛ الأمر الذي يتطلب الإفادة من إمكانات الحاسوب في تحليل هذه البيانات وتمثيلها بالصيغ الصحيحة؛ حتى يتسنى له دراستها دراسة فاحصة وتفسيرها تفسيرًا علميًا.

وعلى الرغم من ذلك فإن الحاسوب لا يستطيع التمييز بين الاستخدام المناسب والاستخدام غير المناسب للأساليب الإحصائية؛ نظرًا لأنه يجري العمليات التي تتطلبها هذه الأساليب بطريقة آلية وفق نظام معين على أي بيانات يدخلها الباحث، فمن خلال البرنامج الإحصائي الحاسوبي مثلًا يمكن إيجاد معامل ارتباط بيرسون بين متغيرين اسميين إذا أعطيناه التعليمات الخاصة بذلك، رغم أن معامل ارتباط بيرسون لا يستخدم مع متغيرين اسميين، كما يمكن الحصول على الوسط الحسابي (Mean) لقيم متوسط متغير اسمي مثل النوع أو الديانة أو الانتماء إلى نادي معين ما دام الباحث قام بترميز أيٍّ من هذه المتغيرات ترميزا كميًا، رغم أن النتائج في هذه الحالة تكون عديمة المعنى؛ أي أن المدخلات غير المناسبة والأوامر غير السليمة في التعامل مع البرامج الإحصائية تؤدي إلى مخرجات أو معطيات خاطئة أو عديمة القيمة ولا تعني شيئا،وهذا يقتضي ألا يحاول الباحث استخدام أسلوب إحصائي معين دون فهم متطلبات هذا الأسلوب وفرضياته وملاءمته لنوع البيانات وطبيعتها ومعنى المعطيات الإحصائية التي تنتج عن استخدامه.

ثانيًا: البرامج الإحصائية الجاهزة

هناك عدد من البرامج والحـزم الإحـصائية المتـوافرة التـي يمكـن أن يـستخدمها الباحـث في تحليـل البيانات، ومن أهم البرامج الإحصائية شائعة الاستخدام في العلوم الاجتماعية والإنـسانية برنـامج حزمـة البرامج الإحصائية للعلوم الاجتماعيـة (Statistical Package for the Social Sciences) والمعروفـة اختـصارًا (SPSS)، وهناك أيضًا حزمة التحليل الإحصائي (MINITAB)، وهي من الحـزم الحاسـوبية الجيـدة التـي يمكن إدارتها في بيئة النوافذ، هناك أيـضًا برنـامج نظـام تحليـل البيانـات (Statistical Analysis System) (SAS) وهو من الحزم الإحصائية البسيطة والقوية التي تفيد البـاحثين في تحليـل البيانـات في المجـالات التطبيقية المختلفة، كما أن برنامج تحليل بيانـات المـسوح (MICROTEST) مـن الـبرامج التـي تفيـد في تحليل البيانات المستمدة من مقاييس الاتجاهات والميول، والمسوح التربويـة والاجتماعيـة والاقتـصادية، كما تهتم بتطويره مؤسسة نظم الحاسوب الأمريكية (National Computer Systems)، بالإضافة إلى ذلك هناك برنامج (Micro Stat)، وهو أحد البرامج الإحصائية الجيدة، ويمتاز بسهولة الاستخدام، ويعتمد على أسلوب الحوار مع المستخدم من خلال طرح الأسئلة والاختيارات المختلفة التي يمكن الاختيـار المناسـب منها.

المطلب الثاني: تعريف بالحزمة الإحصائية للعلوم الاجتماعية

على الرغم من المزايا الفائقـة للـبرامج الإحـصائية الجـاهزة بوجـه عـام، إلا أن الحزمـة الإحـصائية للعلوم الاجتماعية (SPSS) تعتبر أكثر هـذه الـبرامج شـيوعًا عـلى مـستوى العـالم، كـما أنهـا هـي الأكـثر استخدامًا في المجالات البحثية المختلفة في مصـر والعالم العربي. وهذه الحزمة عبارة عـن مجموعـة مـن البرامج المحوسبة، والمعدة مسبقًا (جاهزة) لإدخال ومعالجة البيانات الإحصائية من خلال الحاسوب.

ظهرت أقدم الإصدارات من الحزمة (SPSS) عام ١٩٧٠؛ بهدف التشغيل على الحاسبات الكبيرة، بعد ذلك ظهرت الإصدارات الأولى للحزمة خلال النصف الأول من السبعينيات في القرن العشرين، وكانت تعمل تحت نظام (Dos) مما جعل استخدامها محدودًا؛ وذلك لصعوبة التعامل معها ولاحتياجها إلى أن يكون المستخدم على دراية وخبرة بطريقة وبقواعد كتابه الأوامر اللازمة لتنفيذ عمليات العرض والتحليل.

هذا يعني تدخلًا كبيرًا من المستخدم عند التنفيذ، ولا يستطيع ذلك إلا إذا كان على دراية بطريقة التعامل مع الحاسوب بوجه عام والحزمة الإحصائية بوجه خاص، في وقت لم يكن الحاسوب قد انتشر على نطاق واسع، غير أنه سرعان ما جاء التطور الجوهري في الحزمة الإحصائية للعلوم الاجتماعية (SPSS) منذ بداية التسعينيات من القرن العشرين عندما ظهر الإصدار الخامس والسادس تحت النوافذ باسم (SPSSWIN)، فأصبح التعامل مع الحزمة مسألة سهلة؛ مما أدى إلى انتشار استخدامها، ثم توالت الإصدارات المتطورة من تلك الحزمة.

إن استخدام الحزمة الإحصائية للعلوم الاجتماعية، يقتضي تعرف مكوناتها الأساسية وبيئة عملها حاسوبيا. ومع التسليم باتساع نطاق هذا الموضوع، فإننا سنوضح الأفكار والتطبيقات الأساسية التي تساعد على فهم الحزمة الإحصائية للعلوم الاجتماعية وكيفية استخدامها، وذلك من خلال توضيح النقاط الآتية:

- شريط أدوات الحزمة
- قوائم الحزمة
- الدخول إلى الحزمة
- نوافذ (شاشات) الحزمة، وأهمها: شاشة الإدخال، شاشة المعالجة (الصندوق الحواري)، شاشات التحرير (وهي محرر الجداول، محرر النصوص، محرر الأشكال)، شاشة النسخ، شاشة المخرجات
- الحفظ والإنهاء

أولًا: شريط أدوات الحزمة

يحتوي شريط الأدوات (SPSS Toolbar) على اختصار لبعض العمليات في صورة أيقونات رسومية، كل رسم يمثل عملية معينة مثل: فتح ملف، حفظ ملف... إلخ، ومن الجدير بالذكر أن بعض هذه الأيقونات موحدة الشكل مع التطبيقات الأخرى المختلفة والتي تعمل تحت النوافذ سواء أكانت إحصائية أم غير ذلك مثل (Word, Minitab). ولكل نافذة من نوافذ نظام (SPSS) شريط الأدوات الخاص بها، هذه الأيقونات تساعد المستخدم على تنفيذ الأوامر مباشرة. ويحتوي شريط الأدوات على مجموعة من الأيقونات لكل منها اختصاص معين في عمل الحزمة، مثل:

- فتح ملف
- حفظ الملف
- طبع الملف
- بحث (Find)
- تراجع (Undo)
- إظهار آخر العمليات التي تمت (Dialog Recall)
- إدراج حالة
- إدراج متغير
- تجزئة الملف
- إعطاء أوزان
- اختيار حالة

– اختيار مجموعة حالات

– رؤية المتغيرات (Variables View)

– رؤية البيانات (Data View)

– رؤية رموز قيم الاستجابات (Value Labels)

... إلخ، ويلاحظ إضافة إمكانات جديدة إلى شريط أدوات الحزمة الإحصائية للعلوم الاجتماعية في إصداراتها المختلفة.

ثانيًا: قوائم الحزمة

تعتمد الحزمة الإحصائية للعلوم الاجتماعية (SPSS) على قوائم جاهزة رئيسية (Menu Driven)، وعند النقر على أي قائمة تنسدل قائمة فرعية بالأوامر التي تتضمنها القائمة لأوامر فرعية إلى أسفل (Pull Down)، وبذلك يكون التعامل في الحزمة عن طريق ما يسمى بالقوائم المنسدلة؛ حيث تصاحب كل أيقونة في شريط القوائم قائمة تظهر بمجرد النقر على الأيقونة، توجد داخل كل قائمة مجموعة من الأوامر. وهناك نوعان من الأوامر:

النوع الأول: هو الأمر الذي لا يوجد له أوامر فرعية، ويتم تنفيذه فورًا بالنقر على الأمر مثل: الأمر (Save)، والأمر (Print) من قائمة (File).

النوع الثاني: هو الأمر الذي له أوامر فرعية، بمعنى أنه بمجرد النقر على هذا الأمر، تظهر قائمة بالأوامر الفرعية، والتي يتم اختيار أحدها – حسب العملية المطلوبة.

ويتضمن شريط القوائم لشاشة المحرر مجموعة قوائم مختلفة، أهمها:

١- قائمة ملف (File): وتسمح هذه القائمة بفتح ملف جديد، أو ملف موجود بالفعل، وكذلك من خلال هذه القائمة يمكن حفظ الملفات وطباعتها... إلخ.

٢- قائمة تحرير (Edit): ومن خلال هذه القائمة يمكن إجـراء عمليـات النـسخ (Copy) أو القطـع أو الحذف من الملف.

٣- قائمة عرض (View): تـسمح هـذه القائمـة بإظهـار الأيقونـات، وكـذلك إظهـار وإخفـاء خطـوط مصفوفة البيانات في صفحة محرر البيانات، وتغيير نوع وحجم الخط أو إخفاء دلالات القيم.

٤- قائمة البيانات (Data): تسمح بتعريف متغيرات وتغيـير أسـمائها والقيـام بعمليـات أخـرى مثـل تحويل بيانات، دمج وفرز بيانـات، دمـج ملفـات، إدراج متغـيرات وحـالات، اختيـار متغـيرات وحالات... إلخ.

٥- قائمة التحويلات (Transform): تساعد على إجراء العمليات الحـسابية المختلفـة والتـي تتـضمن استخدام الدوال الرياضية والإحصائية وإعادة التكويد (Recode) وترتيب البيانات.

٦- قائمة العمليات الإحصائية (Analyze): ومن خلالها يتم إجراء العمليات الإحصائية المختلفة.

٧- قائمة الرسومات (Graphics): تستخدم لعمل الرسومات والأشكال البيانية المختلفة.

٨- قائمة الاستخدامات (Utilities): ومن خلالها يحصل المستخدم على معلومات مفصلة عن الملـف والمتغيرات التي يحتويها.

٩- قائمة الإطار (Window): تتحكم هذه القائمة في تكبير وتصغير الـشاشة، وفي التنقـل بـين النوافـذ المفتوحة المختلفة.

١٠- قائمة المساعدة (Help): تتيح هذه القائمة نظامًا كاملًا للمعلومات عن البرنامج، وبمجرد النقر على هذا الأمر يحصل المستخدم على الإجابة التفصيلية عن المعامل الإحصائي الذي يستخدمه.

ثالثًا: الدخول إلى الحزمة

بعد تحميل الحزمة الإحصائية للعلوم الاجتماعية على الحاسوب، فإنه من السهولة الدخول إليها والخروج منها، مثلها مثل أي برنامج جاهز، وهناك طريقتان للدخول إلى الحزمة:

الطريقة الأولى: وهي باستخدام الأمر (Start) الموجود في أسفل شاشة الحاسوب أقصى اليسار، فعند النقر على هذا الأمر تظهر قائمة المحتويات، من بينها (All Programs) وبالنقر عليها تظهر قائمة فرعية بها كل البرامج المتاحة بما فيها (SPSS)، وبالنقر على (SPSS) يفتح البرنامج.

الطريقة الثانية: عمل أيقونة مختصرة (Short Cut Icon) من الحزمة ووضعها على سطح المكتب (الشاشة) أي (Desktop)، ومن خلال النقر المزدوج على تلك الأيقونة يتم الدخول للبرنامج مباشرة.

رابعًا: نوافذ الحزمة

تعمل الحزمة الإحصائية للعلوم الاجتماعية (SPSS) تحت مجموعة من النوافذ (Widows) تمثل في مجموعها منظومة للتعامل مع الحزمة من خلال الحاسوب، ومن أهم نوافذ (شاشات) الحزمة ما يلي:

(أ) شاشة الإدخال:

وتسمى شاشة محرر البيانات (Data Editor Window)، وهذه الشاشة تفتح تلقائيًا عند الدخول إلى البرنامج، ومن خلال هذه الشاشة يمكن عرض محتويات ملف

البيانات الذي سوف يتعامل معه المستخدم، إما لتخزينه أو لإدخال بيانات جديدة أو للعرض أو للتحليل الإحصائي. بيانات هذه الشاشة قابلة للتعديل والحفظ والطباعة. وتتمثل أهم العمليات التي تتم على شاشة المحرر في:

- التحليل الإحصائي للبيانات.
- إنشاء ملف جديد للبيانات.
- فتح وتعديل ملف موجود.
- عمل أية تحويلات على البيانات.
- الحصول على الأشكال البيانية والجداول.
- عرض المعطيات الإحصائية.

وتتكون شاشة المحرر من عدة أجزاء رئيسية، تمكن المستخدم من تسمية الملف، وإدخال البيانات، وحفظ الملف، ووصف المتغيرات، وعرض البيانات المدخلة، وبيان حالة العمل، كما يوجد في هذه الشاشة أيقونات لغلق الشاشة وتصغيرها... إلخ. وفي الجزء العلوي من الشاشة نجد عبارة (Untitled-SPSS Data Editor)، وفيها تتم كتابة اسم الملف، فإذا لم يكتب المستخدم اسم الملف، تظل هذه العبارة موجودة، ومن الطبيعي أن يكتب المستخدم اسم الملف منذ بداية التشغيل، ويظل هذا الاسم ثابتًا للملف.

(ب) شاشة المعالجة (الصندوق الحواري):

الصندوق الحواري (Dialogue Box) عبارة عن تصميم مرسوم يتضمن قسمين، قسم أيسر وآخر أيمن، وفي الجزء الأسفل والأعلى من الصندوق مجموعة من الأوامر الخاصة بالعمليات الإحصائية. وعند طلب إجراء أية عملية إحصائية من الحزمة الجاهزة

(SPSS) فإن البرنامج يظهر صندوقًا حواريًا يحتوي على مجموعة من الخيارات، كل خيار يمثل له أمر معين، ويظهر هذا الأمر أعلى الصندوق بمجرد النقر على هذا الأمر، موضحًا العملية الإحصائية المطلوبة (والتي تم اختيارها).

وفي الجزء الأيسر من الصندوق تظهر قائمة بأسماء جميع المتغيرات الموجودة في الملف النشط والمفتوح على شاشة المحرر، ويمكن التجول في هذه القائمة من أعلى إلى أسفل والعكس عن طريق الفأرة.

أما الجزء الأيمن من الصندوق الحواري، فهو عبارة عن شكل مربع أعلاه كلمة ((Variable(s)) وفي هذا الجزء يتم نقل المتغير أو المتغيرات المطلوبة، بمعنى نقلها من الجزء الأيسر إلى الجزء الأيمن لتجرى عليها العمليات الإحصائية المطلوبة، وذلك من خلال النقر على تلك المتغيرات، فيظهر سهم أسود في الجزء الفاصل بين الجزء الأيمن والجزء الأيسر، ويكون هذا السهم متجهًا نحو الجزء الأيمن، وبالنقر على هذا السهم يتم نقل المتغير المحدد من الجزء الأيسر إلى الجزء الأيمن.

وفي الجزء الأسفل من الصندوق تظهر الأزرار التي توضح كل الأوامر المعنية بالعملية الإحصائية التي حددها المستخدم، ومن خلال إعطاء الأمر المطلوب- بالنقر على هذه الأزرار- ثم النقر على الأمر (Continue) ثم الأمر (OK) الموجودين أعلى يمين الصندوق الحواري يتم تنفيذ العملية الإحصائية المطلوبة، وتظهر المعطيات الإحصائية على شاشة العرض.

وهناك مجموعة من الأزرار أقصى يمين الصندوق، وهي بصفة دائمة موجودة في أي صندوق حواري ولها وظائف ثابتة، من هذه الأزرار:

- (Reset): ومن خلاله يتم إلغاء كل الاختيارات التي تمت في الصندوق الحواري.
- (Cancel): ومن خلاله يتم إلغاء الأوامر التي سبق إعطاؤها، ويتم إغلاق الصندوق الحواري.

- (Paste): ومـن خلالـه يـتم إرسـال الأوامـر الخاصـة بالعمليـة الإحـصائية- إلى نافـذة الأوامـر، وتشغيل هذه الأوامر مرة أخرى.

- (Help): ومن خلاله يمكن معرفة كل المعلومات عـن المعامـل الإحـصائي الـذي يـتم اسـتخدامه بدلالة البيانات الموجودة في الصندوق الحواري.

هـذا بالإضـافة إلى أوامـر فرعيـة أخـرى، تعـرض عـلى المـستخدم إمكانـات إضـافية قد يريدها.

(جـ) شاشة التحرير:

وهي تضم: محرر الجداول، محرر النصوص، محرر الأشكال.

(١) محرر الجداول:

تسمح الحزمـة (SPSS) بتعـديل أيـة جـداول معروضـة في صـفحة النتـائج، وذلـك بواسـطة محـرر الجداول (Pivot Table Editor)، فمن خلاله يتم تعديل الجداول بـأكثر مـن طريقـة منهـا التعـديل مـن خلال محرر الجداول، وذلك بوضع السهم على الجدول المراد تعديله، ثم النقر على الجـزء الأيمـن عـلى الفأرة، فيتم تحديد الجدول بخطوط (إضاءته)، وهذا يعني أن الجدول يستجيب للتعديل. كـما يمكـن باستخدام هذه الإمكانية، تعديل أو إلغاء الجداول أو إلغاء صفوف أو أية أعمدة من الجـدول، ويمكـن أيضًا تغيير الألوان وإنشاء جداول متعددة الاتجاهات.

(٢) محرر النصوص:

يخـتص محـرر الـنص (Text Output Editor) بتعـديل النـصوص التـي تظهـر مـصاحبة للأشكال البيانية مثل: عنوان أو ملحق الشكل أو الجدول. ويـتم التعامـل مـع محـرر الـنص بالنقر المزدوج على الجزء الأيمن للفأرة على النص المراد تعديله فيتم تحديده بخطوط، وبعد

ذلك نجري التعديل المطلوب، يمكن العودة أيضًا على شاشة النتائج بالنقر المزدوج على النص المعدل بعد الانتهاء من التعديل، التعديل يشمل اللون ونوع الخط وحجمه... إلخ.

(٣) محرر الأشكال:

عن طريق محرر الأشكال (Chart Editor) يمكن إجراء كافة التعديلات على الأشكال البيانية والرسومات، ويتم ذلك بعدة طرق منها تعديل اللون وحجم الخط ونوعه وتغيير المحاور. ويتم الوصول إلى هذه الشاشة من صفحة النتائج بالنقر المزدوج بالزر الأيمن للفأرة على الشكل المراد تعديله. ويعمل محرر الأشكال من خلال شريط (Chart Tools)، والذي يتضمن أيقونات مميزة لكل منها استخدام معين؛ بحيث يمكن إجراء أي تعديل على الأشكال المرسومة، مثل: ملء الفراغات في الشكل البياني، التلوين، تغيير أشكال الخطوط أو الأعمدة، إضافة دليل للرسم، تغيير حجم الخط، توصيل النقاط على الخط البياني، فصل شريحة أو أكثر من رسمة الدائرة... إلخ.

(د) شاشة النسخ:

وتسمى شاشة التعليمات (Syntax Editor)، هي الشاشة التي تنسخ عليها الأوامر الخاصة بالعمليات الإحصائية التي حددها المستخدم، ويتم هذا النسخ باستخدام الأمر (Paste)، وتحتوي شاشة التعليمات على صورة من الأوامر التي قام بها المستخدم لعمليات العرض أو التحليل التي استخدمها من قائمة الخيارات. وتكون الأوامر مكتوبة بلغة الحزمة (SPSS) ويمكن تعديلها وحفظها وإعادة تشغيلها.

(هـ) شاشة المخرجات:

شاشة المخرجات (Viewer Window)، هي الشاشة التي تظهر عليها المعطيات الإحصائية، وتفتح هذه الشاشة غيابيًا عند طلب تشغيل أية عملية إحصائية ينتج منها نتائج (جدول أو شكل أو مقياس إحصائي... إلخ). يعرض على هذه الشاشة كل النتائج

الإحصائية (جداول إحصائية، أشكال بيانية، مقاييس إحصائية، اختبارات إحصائية... إلخ)، وهذه المعطيات قابلة للتعديل والحفظ. بمعنى أنه يمكن تعديل الجداول والأشكال البيانية والنصوص وغيرها، كما تسمح الشاشة بحفظ تلك المعطيات، وكذلك بالانتقال منها إلى أية تطبيقات أخرى. وتنقسم شاشة المخرجات إلى جزأين أحدهما لليمين والآخر لليسار. فالجزء الموجود على اليمين مخصص لعرض محتويات الشاشة تفصيليًا من جداول وأشكال بيانية ونتائج وتحليلات إحصائية.. أما الجزء الذي على اليسار، فيحتوي على فهرس تفصيلي لكل محتويات الشاشة، من خلاله يمكن التنقل بسهولة بين محتويات الشاشة...هذا بالإضافة إلى نوافذ أخرى ضمن بيئة عمل الحزمة، والتي تقدم تسهيلات متنوعة، ويلاحظ أن الإصدارات المختلفة من الحزمة تتضمن إمكانات جديدة باستمرار.

خامسًا: الحفظ والطباعة والإنهاء

يوجد بالحزمة الإحصائية (SPSS) نظام حفظ وفتح وطباعة الملفات والخروج من البرنامج، ففيما يخص حفظ ملفات، يتم حفظ محتويات أيه شاشة من الشاشات المذكورة باستخدام الأمر (Save) أو الأمر (Save As)، وهما أمران موجودان في القائمة (File)، الأمر الأول يحفظ الملف مباشرة، ويتم ذلك أوتوماتيكيًا بطريقة قد لا يلاحظها المستخدم، يعطي البرنامج رسالة على شريط الحالة أسفل الشاشة يوضح فيه أن عملية الحفظ جارية أو انتهت، أما الأمر (Save As)، فيتم حفظ الملف باسم جديد يحدده المستخدم، كما يمكن حفظ الملفات في المكان الذي يرغبه المستخدم (Save in). فيما يخص فتح الملفات (Open) فإنه - من خلال هذا الأمر- يمكن فتح الملف المطلوب بعد تحديد هذا الملف، سواء كان يتضمن بيانات، أو نتائج أو رسومات... إلخ، أما فيما يخص طباعة الملفات (Print)، فإنه يتم اختيار هذا الأمر من قائمة (File) وبالنقر على الأمر (Print) تتم طباعة الملف أو بعض أجزائه حسبما يريد المستخدم، ويتم الخروج من البرنامج من خلال علامة الإغلاق، أو من خلال الأمر (Exit).

سادسًا: مزايا الحزمة

إذا كانت الحزمة الإحصائية للعلوم الاجتماعية، تمكننا من إجراء العمليات الإحصائية المعقدة بسرعة وسهولة ووضوح، فإن هناك مزايا أخرى نخص منها بالذكر:

- استخدام ما يسمى بالقوائم (Menu)، ومن خلال النقر على أية أيقونة (Icon) في شريط القوائم، تظهر قائمة بمجموعة الأوامر الفرعية التي يختار المستخدم (User) منها ما يشاء بمجرد النقر دون كتابة الأمر.
- عند اختيار الأمر من القائمة يحصل المستخدم على ما يسمى بالصناديق الحوارية الرئيسية والفرعية، والتي تسهل عملية الاختيار للأوامر دون عناء، وبذلك يسهل تنفيذ أية أمر دون كتابة نص الأمر المطلوب تنفيذه.
- الوصول إلى القوائم المختصرة واستخدامها في أغراض عدة، وذلك باستخدام الفأرة (Mouse)؛ مما يمكّن المستخدم من التفاعل مع البرنامج بسهولة.
- إتاحة عملية السحب (Dragging) والإسقاط (Drop) وذلك للجداول والأشكال البيانية داخل الحزمة (SPSS) وخارجها في برامج وتطبيقات أخرى مثل (Word) أو حزم التطبيقات الإحصائية الأخرى.
- السهولة في تسمية الملفات، مع إمكانية التسمية باللغة العربية.
- إمكانية التعامل مع أكثر من تطبيق في نفس الوقت، وهذا يعني أنه يمكن التعامل مع الحزمة (SPSS) وأي حزم أخرى في نفس الوقت، وذلك بالنقل والنسخ والحذف... إلخ.
- التحسن الكبير والمتواصل في كيفية عرض وتلخيص النتائج والرسومات التي يصل عليها المستخدم (User)، من ذلك على سبيل المثال تحسين التعامل مع البيانات

(Improved Data Access)؛ وذلك لتوفر إمكانية التعامل مع ملفات البيانات الكبيرة دون الحاجة إلى مساحة تخزينية مؤقتة أكبر في ذاكرة الحاسب.

- إمكانية تشغيل الحزمة أكثر من مرة (Multiple Sessions) في فترة التشغيل الواحدة، وبالتالي يمكن التعامل مع أكثر من مجموعة من البيانات والتنقل بين شاشات مختلفة للحزمة من نفس النوع (أكثر من شاشة للمحرر مثلًا أو للنتائج) في نفس الوقت.

- يمكن الدخول مباشرة من الحزمة (SPSS) إلى إكسل (Excel)؛ حيث يمكن قراءة ملفات إكسل مباشرة من الحزمة (SPSS)، كما يمكن قراءة بيانات لمتغيرات ليست من نفس النوع (بعضها رقمي والآخر لفظي) في عمود واحد، كما يمكن تحويل البيانات من الحزمة إلى إكسل والعكس.

خلاصة الفصل الأول

تناول هذا الفصل مدخلًا تعريفيّا بعلم الإحصاء وأهميته للبحث العلمي، وتتلخص موضوعاته الأساسية في:

إن الإحصاء كممارسة نشأ في إطار التنظيم السياسي للدولة القومية في أوروبا منذ الثلث الأخير من القرن الثامن عشر (عام ١٧٧٠)، وسرعان ما تطورت الأصول الرياضية لعلم الإحصاء بفضل علماء الرياضيات الأوروبيين وغيرهم، والذين تواصلت جهودهم العلمية حتى وصل الإحصاء إلى ما هو عليه الآن كعلم متطور يرتبط بكافة مجالات الحياة.

الإحصاء دعامة أساسية لطرق البحث العلمي ذات الطبيعة الكميّة (Quantitative)، والتي تقوم على إجراء دراسات ميدانية وتحليلية، وتجارب وملاحظات ومقابلات؛ حيث يتم الحصول منها على بيانات؛ ومن ثمّ تنظيم تلك البيانات ومعالجتها واستخلاص نتائج موضوعية تصف الموضوع أو الظاهرة، وقد تتعدى نطاق الوصف إلى التفسير وصياغة القوانين والنظريات العلمية.

التصنيف الإحصائي هو وضع الوحدات المتجانسة أو المتشابهة – من حيث صفة معينة أو سلوك معين- في مجموعة واحدة تختلف عن غيرها من المجموعات. والتصنيف قد يكون وصفيّا فقط، بمعنى أنه لمجرد التصنيف دون أي دلالة كميّة، كما قد يكون كميّا، بمعنى أنه يوضح التصنيف ودرجة التمايز.

المتغير (Variable) هو الخاصية أو الصفة أو السلوك محل الدراسة، ومن منظور القياس الإحصائي تتعدد أنماط المتغيرات وفق عدة معايير، فهناك المتغيرات النوعية مقابل

المتغيرات الكميّة. فالمتغيرات النوعية (Qualitative Variables)، هي المتغيرات التي تعبر عن خاصية معينة من حيث وجودها أو عدم وجودها، ولا يمكن ترتيب الأفراد من الأصغر إلى الأكبر، أومن الأكبر إلى الأصغر حسب هذه الخاصية، وإنما يمكن تصنيفهم فقط. أما المتغيرات الكميّة، فهي تلك المتغيرات التي تقاس بمقدار عددي أو كميّة (Quantity)؛ بحيث يمكن ترتيب الأفراد من الأصغر إلى الأكبر، أو من الأقل إلى الأكثر أو العكس. والمتغيرات الكميّة قد تكون متصلة أو منفصلة، فالمتغيرات الكميّة المتصلة هي التي تأخذ قيمة صحيحة أو كسرية، أما المتغيرات الكميّة المنفصلة فهي تلك المتغيرات التي تأخذ قيمًا صحيحة فقط.

إن المتغير المستقل (The Independent Variable) هو المتغير الذي قد يؤثر أو يحتمل أن يؤثر في متغير آخر، أما المتغير التابع (The Dependent Variable) فإنه المتغير الذي يحدث فيه التغير أو يقع عليه التأثير بسبب المتغير المستقل.

من منظور القياس الإحصائي،فإن المتغيرات (Variables) تصنف إلى متغيرات اسمية (Nominal)، ومتغيرات رتبية (Ordinal)، ومتغيرات فئوية أو فترية (Interval)، ومتغيرات نسبية (Ratio).

المتغير الاسمي هو الخاصية أو الصفة من حيث وجودها أو عدم وجودها. أما المتغير الرتبي، فهو المتغير الذي يتضمن معنى المتغير الاسمي، بالإضافة إلى الدلالة الكميّة لأرقام التصنيف؛ أي أن المتغير الرتبي يتضمن تصنيف الأفراد أو الموضوعات في مجموعات متمايزة، بالإضافة إلى إمكانية ترتيبهم تصاعديًا أو تنازليًا.

والمتغير الفئوي أو الفتري هو الذي يعبر عن كمية أو مقدار؛ أي أن له وحدة قياس تعكسها الأعداد من الصفر إلى مالا نهاية، كما أن المتغير الفئوي (الفتري) يتضمن معنى المتغير الرتبي، بالإضافة إلى تساوي الفروق بين الوحدات التي يتم تصنيفها.

أخيرا، فإن المتغير النسبي، هو المتغير الذي تتوفر فيه معاني المتغير الفئوي (الفتري) بجانب كونه يتضمن الصفر المطلق، بمعنى غياب الصفة المقاسة، والمتغيرات النسبية (Ratio Variables) تقع ضمن العلوم الطبيعية، وليس العلوم الاجتماعية بما في ذلك علم الاتصال.

من الضروري أن يكون الباحث على دراية كافية بخصائص المتغيرات؛ لأنه في ضوء هذه الخصائص يتم وضع نظام إدخال البيانات واختيار المعالجة الإحصائية المناسبة، وإذا كانت المتغيرات من منظور القياس الإحصائي تصنف إلى اسمية ورتبية وفئوية (فترية) ونسبية، فإن هناك أنواعًا أخرى من المتغيرات التي يجب أن يعيها الباحث، ومن أهمها: المتغير الدخيل، المتغير الرمزي، المتغير الوسيط.. ولكل منها ظروف استخدامه الخاصة على النحو السابق ذكره.

الإحصاء الوصفي (Descriptive Statistics)، هو كافة المقاييس الإحصائية التي تعتمد على البيانات المستمدة من العينة (Sample) في وصف الظاهرة أو السلوك أو الموضوع أيّا كانت طبيعة هذا الوصف.

أما الإحصاء الاستدلالي (Inferential Statistics)، فيعني التعميم من الجزء على الكل، بمعنى استخدام النتائج المستمدة من دراسة العينة في التعميم على المجتمع.

الإحصاء البارامتري أو الإحصاء المعلمي (Parametric Statistics) يشمل كافة الطرق الإحصائية التي تستند إلى معرفة خصائص المجتمع الذي سحبت منه العينة، ويشترط اعتدالية التوزيع، كما يشترط عشوائية العينة،الاختيار العشوائي للعينة، وأن تتوافر معلومات عن توزيع المجتمع حسب المتغيرات محل الدراسة، وتستخدم طرق الإحصاء البارامتري مع العينات الكبيرة، كما أنها تناسب البيانات الفئوية والبيانات النسبية (Ratio).

أما الإحصاء اللابارمتري (Non-Parametric Statistics)، فهو لا يشترط خصائص معينة في المجتمع الذي سحبت منه العينة، ولا يشترط اعتدالية توزيع الظاهرة في المجتمع، كما لا يشترط الاختيار العشوائي للعينة. وتصلح الطرق اللابارمترية غالبًا للعينات الصغيرة (وإن كانت في بعض الأحيان تستخدم مع العينات الكبيرة)، كما أنها تناسب البيانات الاسمية والرتبية، (وفي بعض الأحيان تستخدم مع البيانات الفئوية (الفترية) والبيانات النسبية).

نستخدم في الإحصاء الكثير من العلامات والرموز والاختصارات والإشارات، والعمليات الجبرية، وعلى الباحث أن يكون على دراية كاملة بها.

أدى التقدم الكبير الذي حدث في الآلات الحاسبة والحواسيب إلى تسهيل وسرعة إجراء العمليات الحسابية مهما كانت معقدة. وهناك اعتقاد خاطئ مفاده أن الحاسوب سيقوم بتحليل البيانات عن طريق البرامج الإحصائية المتوافرة، غير أن الحاسوب لا يستطيع التمييز بين الاستخدام المناسب والاستخدام غير المناسب للأساليب الإحصائية؛ نظرًا لأنه يجري العمليات التي تتطلبها هذه الأساليب بطريقة آلية وفق نظام معين على أي بيانات.

ومن هنا فإن من الضروري أن يكون الباحث على معرفة كافية بطبيعة المتغيرات والمعاملات الإحصائية، ودلالة كل منها، وكيفية تفسيرها، وكيفية تنظيم المعطيات الإحصائية بطريقة صحيحة،وهناك عدد من البرامج والحزم الإحصائية المتوافرة التي يمكن أن يستخدمها الباحث في تحليل البيانات، ومن أهم البرامج الحزمة الإحصائية للعلوم الاجتماعية (SPSS)، حزمة التحليل الإحصائي (MINITAB)، نظام تحليل البيانات (SAS).

مصادر الفصل الأول ومراجعه

(أ) مصادر ومراجع عربية:

رجاء محمود أبو علام (١٩٩٩)، مناهج البحث في العلوم النفسية والتربوية، (القاهرة: دار النشر للجامعات).

سمير كامل عاشور& سامية أبو الفتوح سالم (٢٠٠٣)، العرض والتحليل الإحصائي باستخدام (SPSSWIN)، **الجزء الأول**، (القاهرة: جامعة القاهرة: معهد البحوث والدراسات الإحصائية).

سمير كامل عاشور& سامية أبو الفتوح سالم (٢٠٠٥)، العرض والتحليل الإحصائي باستخدام (SPSSWIN)، **الجزء الثاني**، (القاهرة: جامعة القاهرة: معهد البحوث والدراسات الإحصائية).

صلاح الدين محمود علام (٢٠٠٠)، تحليل بيانات البحوث النفسية والاجتماعية والتربوية، (القاهرة: دار الفكر العربي).

عاطف أحمد منصور(١٩٩٢)، الرياضيات المسلية، (القاهرة: مكتبة ابن سينا).

فتحي عبد اللـه فياض (١٩٩١)، التحليل الإحصائي للبيانات الجغرافية، (القاهرة: مكتبة جامعة عين شمس).

فؤاد أبو حطب & آمال صادق (٢٠١٠)، مناهج البحث وطرق التحليل الإحصائي في العلوم النفسية والاجتماعية والتربوية، (القاهرة: مكتبة الأنجلو المصرية).

(ب) مصادر ومراجع أجنبية:

- Borror, Connie M. Ye, Nong; Mahwah (2003) Statistical Analysis of Normal and Abnormal Data. In: The Handbook of Data Mining. , NJ, US: Lawrence Erlbaum Associates, Publishers, pp. 67-102.

- Denny Burzynski, Wade Ellis (2010) Fundamentals of Mathematics. CONNEXS. Rice University. Houston. TEX.

- Esiner, E. W. (1998) The Enlightened Eye: Qualitative Inquiry and the Enhancement of Educational Practice. Upper Saddle River, NJ: Merrill Prentice – Hall.

- Murism, R.D. (2005) Statistical Modeling in the Sciences. Australia. Pearson Education.

- N.A.I (2004) Computational Statistics and Data Analysis. Special Issue: The Fuzzy Approach to Statistical Analysis. Journal of Multivariate Analysis, Vol. 91(2), pp. 306.

- Shank, G. D. (2002) Qualitative Research: A Personal Skills Approach: NJ: Merrill Prentice – Hall.

- Shepherd, Martin (1985) EMDISP: A Visual Display System with Digital and Analogue Sampling; Behavior Research Methods, Instruments & Computers, Vol. 16(3), pp. 297-302.

- SPSS Manuals. Different Versions. London. Sage.

-Susan Dean. Barbara Illowsky (2009) Collaborative Statistics: Symbols and Their Meanings. The Connexions Project. Creative Commons Attribution. Available at: www.scribd.com/Collaborative-Statistics-Symbols.

- www.sussex.ac.uk/Users/grahamh/ Stats Symbols. A Brief Guide to Some Commonly Used Statistical Symbols.

الفصل الثاني

إعداد البيانات والمعالجة الإحصائية الوصفية

تمهـــيد

بعد الانتهاء من جمع بيانـات الدراسـة ومراجعتهـا، تـأتي الخطـوة اللاحقـة ممثلـة في إدخـال تلـك البيانات وتحليلها، ومن الثابت أن التحليل السليم لبيانات البحوث العلمية يتطلب إعداد تلك البيانات بطريقة معينة؛ بحيث يمكن معالجتها إحصائيّا بما يتفق والإجابة على التساؤلات التـي تـسعى الدراسـة للإجابة عليها أو التحقق من الفروض التي تستهدف الدراسة التحقق منها.

الفصل الحالي يعرض أساسيات إدخـال البيانـات وإعـدادها للتحليـل الإحـصائي مـن منظـور الفكـر والمهارة؛ حيث يتضمن ثلاثة مباحث، يختص المبحـث الأول بتوضـيح عمليـة تجهيـز البيانـات للتحليـل الإحصائي، بمعنى كيفية إعدادها بطريقة صحيحة؛ بحيث يمكن معالجتها إحصائيّا معالجة سليمة تقابـل أهداف البحث؛ وذلك من حيث المراجعة، الإدخال، التنظيم. أما المبحث الثاني، فيناقش أساليب الوصف التكراري والبياني للمعطيات الإحصائية الأولية من خلال التكرارات والنسب والأشـكال البيانيـة، في حـين يأتي المبحث الثالث متضمنًا التعريف بمقاييس النـزعة المركزية للبيانات، ومقـاييس التـشتت، باعتبارهـا من أهم الأساليب الإحصائية شائعة الاستخدام في العلوم الاجتماعية.

المبحث الأول

تجهيز البيانات

يقصد بتجهيز البيانات إعدادها بالطريقة المناسبة؛ بحيث يمكن معالجتها إحصائيًا معالجة سليمة تقابل أهداف البحث، وتتمثل الإجراءات الأساسية لتجهيز البيانات في: المراجعة، الإدخال، التنظيم؛ ومن ثمّ وصف البيانات على النحو الآتي:

أولًا: مراجعة البيانات

قبل إدخال البيانات التي ستتم معالجتها إحصائيًا تتم مراجعتها بدقة متناهية، سواء كانت بيانات تحليلية (مستمدة من تحليل الرسالة الإعلامية)، أو كانت بيانات ميدانية (مستمدة من الجمهور أو القائم بالاتصال أو غير ذلك من الأشخاص الطبيعيين والاعتباريين)، والهدف الأساسي من المراجعة المدققة هو التأكد من أن البيانات صحيحة ومستوفاة، وبموجب هذه المراجعة يتم حذف الاستمارات أو البيانات غير المكتملة، أو المتضاربة، أو غير الدقيقة / غير المنطقية، أو التي تكشف عن مواصفات غير متطابقة (كأن تكون الدراسة على مبحوثين من فئة عمرية معينة ثم نكتشف في المراجعة أن هناك بعض استمارات لمبحوثين من فئات عمرية أخرى، أو أن يكون المخطط أن تتوزع العينة توزيعًا معينًا حسب الخصائص الديموجرافية،ثم نكتشف زيادة أو نقصًا في عدد الاستبانات بما يخل بهذا التوزيع).

وبعد استبعاد الاستمارات غير الصالحة، فإن على الباحث اتخاذ القرار: إما باستكمال العدد المطلوب من الاستمارات بالمواصفات المطلوبة، أو الاكتفاء بما هو موجود،ويتوقف ذلك على حجم العينة وطبيعة البحث وإمكانية الاستكمال من عدمها. فقد يقرر الباحث أن يستكمل الاستمارات غير المستوفاة /غير الصحيحة، وقد يكتفي بالاستمارات الصحيحة، خاصة إذا كان حجم العينة كبيرًا، وكانت البيانات الناقصة قليلة أو غير

أساسية في الدراسة، أو أن عدد الاستمارات المستبعدة قليل ولا يؤثر في حجم العينة ولا يعرقل المعالجة الإحصائية المطلوبة .. وفي كل الأحوال، فإن على الباحث أن يذكر ذلك في تقرير البحث.

ثانيًا: إدخال البيانات

من المسلم به أنه قبل تطبيق أداة جمع البيانات يتم التدقيق فيها من المنظور الإحصائي، فالاستبيان مثلًا، يتم التأكد من أن الأسئلة /البنود بأرقام متسلسلة، وأن الاستجابات لها قيم كميّة (Quantitative Values) واضحة، وأن يتخذ الباحث قرارًا بشأن الأسئلة/ البنود المفتوحة. وبعد الانتهاء من جمع وتسجيل البيانات ومراجعتها على النحو المشار إليه، يتم ترقيم جميع الاستمارات ترقيمًا متسلسلًا، فإذا كان لدينا استبيان مثلًا تم تطبيقه على عينة قوامها ٤٠٠ مفردة، فإن الباحث يقوم بترقيم الاستبانات من (١) إلى (٤٠٠). ويتم وضع النظام الإحصائي لإدخال البيانات باستخدام أحد البرامج الإحصائية الجاهزة؛ وبحيث يتوافق هذا النظام مع مكونات أداة جمع البيانات، ففي حالة الاستبيان مثلًا لا بد أن يكون نظام إدخال البيانات يعكس مكونات الاستبيان؛ من حيث عدد الأسئلة/ البنود والمتغيرات.

وقد ظهرت طرق حديثة في إدخال البيانات باستخدام برامج حاسوبية تقوم بعمل مسح (Scanning) للاستبانات، ومن خلال هذا المسح يتم إدخال البيانات آليّا، وهذه الطريقة تستخدم على نطاق واسع في الكثير من استطلاعات الرأي العام، ومع التأكيد على أهمية طريقة المسح (Scanning)، إلا أن الممارسة الفعلية كشفت عن أنها تنطوي على أخطاء حتى في حالة إعداد الاستبيان وتكويده بالأسلوب الذي تقتضيه هذه الطريقة، وفي تصويب هذه الأخطاء قد يستغرق الأمر وقتًا وجهدًا لا يقل عما لوتم إدخال البيانات بالطريقة التقليدية، على أن طريقة إدخال البيانات بطريقة المسح الآلي يكون لها ميزات في البحوث التي تتعدد فيها أدوات جمع البيانات، وتجري على عينات كبيرة. وهناك الكثير من المراجع المتخصصة بل والتطبيقات المحوسبة في هذا الشأن يمكن للباحثين الرجوع إليها.

ما يعنينا في هذه الجزئية توضيح كيفية إدخال البيانات في الحاسب الآلي بالطريقة التقليدية.فبموجب أرقام الأسئلة/ البنود التي يضمها الاستبيان، يتم وضع النظام الإحصائي لإدخال البيانات فإذا كان السؤال أو البند من النوعية التي يختار عليها المبحوث استجابة واحدة (One-Response)، فإنه يخصص له عمود واحد (Column) في نظام الإدخال، ويكون هذا العمود يحمل رقم نفس السؤال، ويتم إدخال رقم الاستجابة التي اختارها المبحوث على السؤال. لنفرض أن السؤال هو: «هل تستخدم الإنترنت؟» وكانت الاستجابات على هذا السؤال هي:

(١) لا (٢) أحيانًا (٣) دائمًا

فإذا اختار المبحوث الاستجابة الأولى، فإننا ندون رقم ٣ في العمود المخصص لهذا السؤال، أما إذا اختار المبحوث الاستجابة الثانية، فإننا ندون الرقم ٢ ، وإذا اختار المبحوث الاستجابة الثالثة، فإننا ندون الرقم ١ ... وهكذا في جميع الأسئلة ذات الاختيار الواحد. أما إذا كان السؤال من النوعية التي يمكن للمبحوث أن يختار أكثر من استجابة (Multi-Response)، فإنه يخصص له عدد من الأعمدة يساوي عدد الاستجابات أو البدائل. لنفرض أن الاستبيان يتضمن سؤالًا هو: « لماذا تشاهد التليفزيون؟»، وأن هناك عشرة أسباب يختار منها المبحوث، في هذه الحالة لا بد أن يكون لهذا السؤال عشرة أعمدة، فإذا كان السؤال المعني هو السؤال رقم ٧ في الاستبيان مثلًا، فإن نظام الإدخال لهذا السؤال يكون كالآتي:

Q7.1-Q7.2- Q7.3- Q7.4- Q7.5- Q7.6-Q7.7- Q7.8- Q7.9- Q7.10

ويتم إدخال الرقم ١ ليدل على البديل الذي اختاره المبحوث، بينما يتم إدخال الرقم صفر ليدل على البديل الذي لم يختاره المبحوث.... وهكذا في جميع الأسئلة متعددة الاختيارات.

أما فيما يخص الأسئلة المفتوحة (Open Ended Questions)، فإن إدخالها يتطلب الكثير مـن الوقـت والتركيز الذهني. لنفرض أن السؤال هو:«ما هي القنوات التليفزيونية التي تفضل مـشاهدتها»؟، وهـذا السؤال مفتوح؛ أي أن المطلوب من المبحـوث أن يـذكر القنـوات التليفزيونيـة التـي يفـضلها، وبطبيعـة الحال، فإن المبحوث قد يذكر عدة قنوات، وقـد يـذكر قنـاة واحـدة، أو قنـاتين... إلـخ. إن وضـع نظـام إحصائي لإدخال بيانات مثل هذا السؤال يأخذ هذه الحقيقة بعين الاعتبار؛ بحيـث يـتم تخـصيص عـدد كاف من الأعمدة (Columns)، فقد تصل القنوات التي تفضلها العينة إلى مئات القنـوات (لاحـظ أننـا نتحدث عن كل العينة وليس عن الفرد الواحد)، غير أنه من الصعب أن نجد مبحوثًـا يفـضل مـشاهدة أكثر من عشر قنوات، ومن هنا يمكن أن نخصص عشـرة أعمدة لهذا السؤال.

إن هناك عدة طرق لتكويد مثل هذا السؤال وغيره من الأسئلة المفتوحة، أبسط هـذه الطـرق هـي أنه أثناء إدخال البيانات يقوم الباحث بكتابة اسـم القنـاة/ القنـوات التـي اختارهـا المبحـوث في ورقـة مستقلة، ويعطي رقمًا معينًا لكل قناة؛ أي أن هذه الورقة تتـضمن اسـم القنـاة /أسـماء القنـوات التـي اختارها المبحوث، وقرين اسم القناة يوجد رقم معين، لنفرض أن أحد المبحوثين أفاد بأنـه يفـضل: قنـاة الجزيرة، قناة العربية، قناة الأخبار المصرية، قناة MBC. هنا يقوم مدخل البيانات بكتابة هذه القنـوات في ورقة مستقلة، مع إعطاء رقم (١) لقنـاة الجزيـرة، رقـم (٢) لقنـاة العربيـة، رقـم (٣) لقنـاة الأخبـار المصرية، رقم (٤) لقناة MBC.

وفي عملية إدخال البيانات يقوم الباحث بإدخال الرقم (١) ليـدل عـلى قنـاة الجزيـرة، ثـم إدخـال الرقم (٢) ليدل على قناة العربية، ثم الرقم (٣) ليدل على قناة الأخبار المصرية، ثـم إدخـال الـرقم (٤) ليدل على قناة MBC، ويتم الالتزام بذلك في إدخال بيانات هذا السؤال في جميع الاستبانات، بمعنـى أن أي مبحوث يذكر أنه يفضل قناة الجزيرة، يتم إدخال الرقم (١) على أساس أنه تم تكويد قنـاة الجزيـرة بهذا الرقم، وبالتالي يكون الرقم (١) هو الرمز الكودي لقناة الجزيرة.

وهكذا فيما يخص جميع القنوات، فلكل منها رقم كودي مميز، وبالتأكيد فإن القائم على إدخال البيانات سيجد مبحوثين يفضلون مشاهدة قنوات أخرى، وهنا يتم تخصيص رقم كودي جديد لكل قناة جديدة يذكرها المبحوثون، ويقوم مدخل البيانات بتدوين اسم القناة ورقمها الكودي في الورقة المستقلة التي تضم أسماء وأرقام القنوات التليفزيونية. هذه الورقة ما هي إلا دليل التكويد (Codebook) وهي تصبح أساسية ولا غنى عنها أبدًا للقائم على المعالجة الإحصائية، فهذه المعالجة تسفر عن أرقام، فإذا وجدنا أن تكرارات الرقم (١) هي على سبيل المثال ٥٤٢ فإن ذلك يعني أن هناك ٥٤٢ مفحوصًا أفادوا بأنهم يفضلون مشاهدة قناة الجزيرة، فإذا كان حجم العينة ١٠٠٠ مفردة، يمكننا القول إن ٥٤.٢% من العينة أفادوا بأنهم يفضلون قناة الجزيرة، وهكذا بالنسبة لجميع القنوات.

هكذا تستمر عملية إدخال البيانات لجميع الاستبانات، وبعد الانتهاء منها، تخضع البيانات المدخلة للمراجعة مرة أخرى؛ بهدف التأكد من أن البيانات التي تم إدخالها هي بالضبط البيانات الموجودة في الاستبانات، ويمكن أن يتم ذلك على عينة عشوائية (وليس على جميع الاستبانات)، فلكل استبانة رقم متسلسل مدون على الغلاف، وتم إدخالها في النظام الإحصائي بهذا الرقم، ويمكن للباحث بسهولة اختيار أي عدد من الاستبانات بحيث يضاهي بين البيانات التي أدلى بها المبحوثون في هذه الاستبانات وبين البيانات التي تم إدخالها، ومن السهل جدّا اكتشاف أي بيانات خاطئة تم إدخالها؛ ومن ثمّ يتم تصحيحها في النظام الإحصائي.

وهذا الإجراء يدخل ضمن ما يعرف بتنظيف البيانات (Data Cleaning)؛ حيث يتم التأكد من دقة البيانات التي تم إدخالها، وأنها مطابقة ليس فقط لما هو موجود في الاستبانات، وإنما أيضًا مطابقة للمنطق، فإذا كانت قيمة استجابات أحد الأسئلة أو البنود كما هو في الاستبيان هي (١) أو (٢)، فإنها في البيانات المدخلة لا بد أن تكون كذلك (١) أو (٢) فقط (وليس أي رقم آخر)، فإذا تبين وجود أرقام مختلفة يتم الرجوع إلى الاستبانة الأصلية لمعرفة الرقم المدون بها وإدخاله في الحاسوب.

ثالثًا: تنظيم البيانات

ويعني تكييف البيانات إحصائيّا Statistically Adjusting The Data بما يناسب أهداف البحث وفروضه وتساؤلاته، ومن إجراءات التكييف الإحصائي للبيانات نخص بالذكر:

(أ) **إعادة التكويد:** قد يقتضي الأمر إعادة التكويد Recoding لبعض الاستجابات، بمعنى تغيير أرقام القيم الكمية للاستجابات بحيث تعكس شدة هذه الاستجابات، وقد يتم ذلك أثناء إدخال البيانات، كما قد يتم بعد إدخال البيانات. افترض أن الاستجابات على سؤال: «هل تشاهد التليفزيون؟» كانت كالآتي:

(١) بانتظام (٢) غالبًا (٣) نادرًا

لقد تم إدخال القيم الكمية لهذه الاستجابات كما هي، ومن الملاحظ أن الاستجابة الأولى (بانتظام) قد أخذت الرقم الكودي (١). أما الاستجابة (أحيانًا) فقد أخذت الرقم الكودي (٢)، وأخيرًا فإن الاستجابة (نادرًا) أخذت الرقم الكودي (٣)، لقد تم إدخال أرقام الاستجابات في النظام الإحصائي على هذا النحو. لكن المنطق يقتضي أن الاستجابة الأولى (بانتظام) تأخذ قيمة كميّة أعلى من الاستجابة (أحيانًا) ، وأن هذه الاستجابة (أحيانًا) بدورها تأخذ قيمة كميّة أعلى من الاستجابة الأقل (نادرًا)، ومن هنا تتم إعادة تكويد الاستجابات بحيث تكون :

(٣) بانتظام (٢) غالبًا (١) نادرًا

وبموجب ذلك أصبحت القيم الكميّة للاستجابات متفقة مع شدتها، فالاستجابة الأشد أخذت قيمة أعلى، والاستجابة الأقل شدة أخذت قيمة أقل، والاستجابة الأضعف أخذت قيمة كميّة أقل من قيمة كلّ من الاستجابتين. إن عملية إعادة التكويد تغطي مدى واسعًا جدًا من الاستخدامات، فقد يرى الباحث مثلًا ضرورة إعادة تكويد الاستجابات

على بنود المقاييس الفرعية؛ بحيث تكون ذات دلالة معينة (كأن تكون الدرجة المرتفعة على هذه المقاييس ذات مغزى مختلف عن الدرجة المنخفضة، وقد يجد ضرورة إعادة تكويد استجابات بنود أو أسئلة معينة بما يعكس حقيقة العلاقات بينها... إلخ).

(ب) **تكوين المتغيرات:** بمعنى تكوين متغيرات جديدة من البيانات الموجودة التي تم إدخالها، فإذا كان لدينا استبيان يضم عشرة بنود تقيس معدل استخدام وسائل الإعلام، وعشرة بنود أخرى لقياس المعرفة بالقضايا السياسية، وعشرين بندًا لقياس المستوى الاقتصادي الاجتماعي، فإن مجمل المقياس يكون لدينا أربعون بندًا ومتغيرًا. في هذه الحالة يتم تكوين ثلاثة متغيرات، الأول هو معدل استخدام وسائل الإعلام (ويتم تكوينه من استجابات المبحوثين على البنود العشرة التي تقيس هذا المعدل)، المتغير الثاني هو المعرفة السياسية (ويتم تكوينه من استجابات المبحوثين على البنود العشرة التي تقيس هذه المعرفة)، المتغير الثالث هو المستوى الاقتصادي الاجتماعي (ويتم تكوينه من العشرين بندًا التي تقيس المستوى الاقتصادي الاجتماعي). هكذا أصبح لدينا ثلاثة متغيرات جديدة تم تكوينها من البيانات الموجودة، وهذه المتغيرات هي متغيرات تجميعية لمقاييس فرعية.

(جـ) **تحديد المستويات:** بمعنى التعرف على مستوى مجمل الاستجابات على بنود تقيس متغير معين (كمجموعة بنود تقيس دوافع استخدام الإنترنت، أو مجموعة بنود تقيس المعرفة السياسية... إلخ)، فقد يتم تحديد المستوى مثلًا على أنه (منخفض، متوسط، مرتفع)، ويستند تحديد المستوى إلى الوسط الحسابي والانحراف المعياري للقيم الكميّة التي تعكس استجابات المبحوثين، فالمبحوثون الذين حصلوا على درجة تعادل قيمة الوسط الحسابي للعينة ناقص واحد انحراف معياري يتم تصنيفهم في المستوى المنخفض، أما المبحوثون الذين حصلوا على درجة تعادل ما بعد درجة ذوي المستوى المنخفض مرورًا بقيمة الوسط الحسابي للعينة

(Mean) + واحد انحراف معياري، فيصنفون ضمن المستوى المتوسط (Medium Level)، بينما المبحوثون الذين حصلوا على درجة أعلى من ذلك، يتم تصنيفهم ضمن المستوى المرتفع. كمثال على ذلك نفرض أنه في دراسة عن المعرفة السياسية تبين أن متوسط درجة العينة ١٨ بانحراف معياري ٣ في هذه الحالة يتم تحديد المستويات كالآتي:

المستوى المنخفض = ١٥

المستوى المتوسط (من ١٦ إلى ٢١)

المستوى المرتفع (٢٢ فأكثر)

ولهذه الطريقة ميزة أنها تجعل توزيع العينة اعتداليّا؛ وتؤدي إلى معطيات إحصائية سليمة للعلاقات أو الفروق لأنها تتلافى عيوب التصنيف القائم على تساوي أقسام درجة المقياس؛ حيث لوحظ أن بعض الباحثين إذا كانت الدرجة على المقياس تبلغ ٣٠ درجة، فإنه يضع تصورًا مفاده أن المبحوثين الذين حصلوا على (١٠) درجات فأقل يتم تصنيفهم ضمن المستوى المنخفض، أما المبحوثون الذين حصلوا على (١١ إلى ٢٠ درجة) فإنهم يصنفون ضمن المستوى المتوسط ، بينما المبحوثون الذين حصلوا على (٢١ إلى ٣٠ درجة) يتم تصنيفهم ضمن المستوى المرتفع. هذا التصنيف يقوم على افتراض خاطئ هو تساوي أو تقارب عدد المبحوثين في كل مستوى لدرجة المقياس (الثلث الأول، الثلث الثاني، الثلث الثالث) رغم أن معظم التكرارات قد تتركز في مستوى واحد (المستوى المرتفع مثلًا) وتنخفض بشدة أو قد تتلاشي تمامًا في أحد المستويات، كما أن تصنيف المبحوثين على هذا النحو يخل بأسس التوزيع الاعتدالي.

(د) التحويل: كثيرًا ما تتطلب البحوث تحويل البيانات، **و**تتعدد أشكال التحويل، ومن أهمها تحويل الدرجات الخام (Raw Scores) إلى درجات معيارية

(Standard Scores)، سواء في صورة درجات زائية (Z Score) أو في صورة درجات تائية (T Scores)، ففيما يخص الدرجات الزائية والمعروفة باصطلاح (Z Score)، فإنها درجات معيارية لتوزيع معين متوسطه الحسابي صفر وانحرافه المعياري واحد صحيح، هذا التوزيع، قد يتمثل مثلًا في توزيع عينة من المبحوثين حسب الوقت اليومي المنقضي في استخدام الإنترنت أو غير ذلك من التوزيعات.

والدرجات الزائية أو (Z Score) تستخدم كبديل عن الدرجات الخام التي لا تصلح لمقارنة درجات على مقاييس مختلفة، أو مقارنة أفراد ينتمون إلى مجتمعات مختلفة، فمتوسط الدرجة على مقياس المستوى الاقتصادي الاجتماعي لا يصلح مقارنته بمتوسط الدرجة على مقياس استخدام وسائل الإعلام؛ لأن وحدات القياس مختلفة تمامًا في المقياسين، كما أن متوسط درجة مجموعة من النخبة المثقفة على مقياس قراءة الصحف لا يصلح للمقارنة بمتوسط درجة مجموعة من ذوي التعليم الابتدائي على المقياس نفسه؛ إذ إن متوسط درجة مجموعة النخبة يكون أعلى بكثير من متوسط درجة مجموعة ذوي التعليم الابتدائي.

وقد يكون لدى الباحث مقياسان: الأول مثلًا ، يقيس المعرفة السياسية وتتراوح الدرجة عليه ما بين (٣٠) إلى (٨٠) درجة، أما المقياس الثاني فيقيس معدل استخدام الصحف وتتراوح الدرجة عليه مابين (٥) إلى (١٠) درجات، فلا يكون من المناسب مقارنة متوسطات الدرجات الخام على المقياس الأول بالدرجات الخام على المقياس الثاني، ومن هنا تستخدم الدرجات المعيارية للمقارنة (سوف نتناول ذلك بالتفصيل ضمن الدرجات المعيارية كمقياس للتشتت).

بالإضافة إلى الدرجات المعيارية، هناك أيضًا التحويلات الخطية (Linear Transformations)، وهذه التحويلات تعني إدماج أنماط مختلفة من البيانات في نمط معياري (أي موحد)، لنفرض أن دراسة تهتم بالعلاقة بين التعرض لوسائل الإعلام

ومستوى الانفتاح على العالم. إن مؤشرات الانفتاح على العالم قد تتمثل في عدة متغيرات مثل: مشاهدة الدراما الأجنبية، إتقان اللغات الأجنبية، الاهتمام بقضايا المجتمعات الأخرى، قـراءة الـصحف الأجنبيـة، المشاركة في المؤتمرات والندوات الدولية، السفر للخارج، الاشتراك في عضوية جمعيـات دوليـة والنـشاط فيها... هنا نلاحظ أن هذه المتغيرات لكل منها وحدة قياس مختلفة، وبالتالي يستخدم الباحث التحويلة الخطية للخروج بمقياس معياري لتحديـد (مـستوى الانفتـاح عـلى العـالم)، وللتحـويلات الخطيـة صـيغ رياضية ثابتة،وتستخدم في الكثير من المعاملات الإحصائية، مثل تحليـل الانحـدار المتعـدد، وفي التحليـل العاملي، وتحليل التباين المتعدد.

المبحث الثاني

الوصف التكراري والبياني

بموجب الإجراءات السابقة يصبح بإمكان الباحث التعـرف عـلى طبيعـة المعطيـات الإحـصائية مـن خلال الوصف الأولي؛ أي أننا هنا بصدد تحليل إحصائي مبسط، تتحدد في ضـوئه المعـاملات الإحـصائية التي يمكن استخدامها بما يتفق والإجابـة عـلى التـساؤلات والتحقـق مـن الفـروض. وتتـضمن المعالجـة الوصـفية معرفـة التكـرارات والنـسب (Frequencies and Percentages) وخـصائص النزعـة المركزيـة والتشتت؛ ومن ثمّ تبسيط البيانات وسهولة تلخيصها في صورة موجزة توضح أهم خواصها؛ بحيث يصل الباحـث إلى وصـف موضـوعي للمتغـيرات وتنظيمهـا في جـداول ورسـوم بيانيـة، ومنحنيـات وأشـكال توضيحية تبين معالم المتغيرات وخصائص البيانات.

ومن خلال الوصـف الإحـصائي يـتم التمهيـد للتحليـل الإحـصائي المناسـب؛ لأن الوصف الإحصائي يوضح الخواص الإحصائية للمتغيرات؛ ومـن ثـمّ- بنـاءً عـلى هـذا

الوصف- يتم تحديد نوع التحليل الإحصائي ومستواه، وبإجراء الوصف الإحصائي المناسب يتم الحصول على المعطيات الكميّة (Quantitative Outcomes)، ثم تنظيم تلك المعطيات وعرضها في جداول واضحة، وإجراء المعالجات الإحصائية اللازمة بما يقابل أهداف البحث. ويتخذ الوصف الإحصائي أشكالًا متعددة أهمها التكرارات والنسب، والتي يمكن التعبير عنها بالأشكال البيانية المختلفة.

أولًا: التوزيعات التكرارية

وهي التكرارات والنسب المئوية (Frequencies and Percentages) للمتغيرات أيّا كانت، ويتم التعبير عنها في جداول منظمة، وعلى الرغم من أن التكرارات والنسب أبسط مستويات الوصف الإحصائي، إلا أنها أساس كافة أشكال ومستويات المعالجة الإحصائية، ومن خلال التكرارات والنسب، يمكن الوقوف على الخصائص الأساسية لتوزيع المفردات حسب جميع المتغيرات، ففي بحوث تحليل المحتوى تستخدم الجداول التكرارية للتعبير عن تكرارات ونسب كل فئة من فئات استمارة التحليل، فإذا كان هناك فئة عن مصادر الأخبار الصحفية أو الإذاعية مثلًا، وكان عدد الأخبار التي تم تحليلها (٢٠٠) خبر، فإنه بعد تفريغ هذه الأخبار وتصنيفها حسب المصدر، يتم تنظيمها في جدول تكراري بسيط كالآتي:

المصادر	ك	%
وكالات أنباء	١٦٨	٨٤
مراسل	٢١	١٠.٥
الإنترنت	١١	٥.٥
المجموع	٢٠٠	١٠٠

فهذا الجدول يلخص توزيع عينة الأخبار محل الدراسة حسب مصادرها، ومنه يتضح أن ٨٤ % من الأخبار تم الحصول عليها من وكالات أنباء، مقابل ١٠.٥% من المراسلين وكذلك ٥.٥% من الإنترنت. وفي البحوث الميدانية - حيث يتم تطبيق الاستبيانات أو الاختبارات والمقاييس على عينة من المبحوثين- يتم استخراج التكرارات والنسب التي تعكس أنماط الاستجابات على البنود/التساؤلات التي تشكل الاستبيان أو الاختبار أو المقياس، ومن ثمّ التعبير عن ذلك في جداول تكرارية واضحة، فإذا كانت العينة تضم ٤٠٠ مفردة، فإن التكرارات والنسب تعني توزيع هذه العينة حسب الخصائص الديموجرافية وأنماط الاستجابة على بنود وأسئلة الاستبيان، فالباحث مثلًا يعبر عن خصائص العينة من حيث: النوع، السن، التعليم، محل الإقامة وغير ذلك من المتغيرات من خلال جدول تكراري يتضمن تكرارات ونسب مفردات العينة موزعة حسب كل الخصائص الديموجرافية، كما أن الجداول التكرارية تتضمن توزيع العينة حسب أنماط الاستجابات على كافة البنود والأسئلة، فمعدل قراءة الصحف اليومية مثلًا يتم التعبير عنه بالتكرارات والنسب في جدول واضح كالآتي:

معدل القراءة	ك	%
يوميّا	١٢٥	٣١.٣
من ٥ إلى ٦ أيام	١٣٨	٣٤.٥
من ٣ إلى ٤ أيام	٥٢	١٣
يوم أو يومان	٨٥	٢١.٢
المجموع	٤٠٠	٤٠٠

فهذا الجدول يعبر بوضوح وبساطة عن توزيع عينة كبيرة حسب معدل قراءة الصحف، ولنا أن نتصور حجم البيانات وعدم معناها دون هذا التنظيم.

أنماط الجداول التكرارية

الجداول التكرارية قد تكون بسيطة، كما قد تكون مزدوجة (Cross-Tabs)، وقد تكون متعددة المتغيرات (Multivariate Variables Tabs) ، فالجدول البسيط هو الجدول الذي يتضمن تكرارات متغير واحد، كما هو الحال في الجدولين السابقين (فالجدول الأول يوضح توزيع الأخبار حسب متغير واحد فقط هو مصدر الخبر، كما أن الجدول الثاني يتضمن توزيع العينة حسب متغير واحد فقط هو معدل قراءة الصحف اليومية)، أما الجدول المزدوج فهو الجدول الذي يتضمن التوزيع حسب متغيرين معًا، مثال ذلك توزيع عينة الأخبار حسب المصدر والمدة الزمنية، أو توزيع عينة المبحوثين حسب النوع (ذكور & إناث) ومعدل قراءة الصحف اليومية؛ حيث يمكن التعبير عن هذين المتغيرين في جدول مزدوج من خلال تقسيم عمود التكرارات إلى عمودين (Two Columns) أحدهما للذكور، والآخر للإناث، ويتخذ الجدول الشكل التالي:

معدل قراءة الصحف	النوع	
	ذكور	إناث
يوميّا	٤٥	٨٠
من ٥ إلى ٦ أيام	٨٠	٥٨
من ٣ إلى ٤ أيام	٣٠	٢٢
يوم أو يومان	٤٥	٤٠
المجموع	٢٠٠	٢٠٠

ويمكن توزيع التكرارات حسب نسبتها المئوية سواء على مستوى الذكور فقط أو الإناث فقط أو على مستوى العينة ككل، ويتم تكوين الجدول المزدوج وتفريغ البيانات فيه يدويّا (في حالة العينات الصغيرة)، وذلك بفحص كل استبيان على حدة ووضع علامة في الخانة التي تعبر عن الحالة، فالحالة الأولى مثلًا قد تكون لمبحوث من الذكور، ويقرأ الصحف يوميّا، وهنا يتم وضع علامة في الخانة الخاصة بالذكور الذين يقرؤون الصحف

يوميًا.. وهكذا يستمر تفريغ الاستبيانات إلى أن تنتهي، ويكون لدى الباحث عدد من الجداول المزدوجة حسب عدد المتغيرات مجال الدراسة.

أما الجداول متعددة المتغيرات (Multivariate Variables Tabs)، فهي الجداول التي تتضمن توزيع التكرارات حسب أكثر من متغيرين (ثلاثة متغيرات أو أكثر)، فقد يتطلب البحث توزيع المبحوثين حسب معدل قراءة الصحف مصنفين وفق النوع (ذكور & إناث)، والسن (أقل من ٣٠ سنة & ٣٠ سنة فأكثر)، وفي هذه الحالة يتم تصنيف الذكور إلى مجموعتين حسب السن (أقل من ٣٠ سنة & ٣٠ سنة فأكثر)، وتصنيف الإناث أيضًا حسب السن (أقل من ٣٠ سنة & ٣٠ سنة فأكثر)، ويتم توزيع تكرارات معدل قراءة الصحف لكلا الجنسين في جدول واضح على النحو التالي:

معدل قراءة الصحف	ذكور		إناث	
	أقل من ٣٠ سنة	٣٠ سنة فأكثر	أقل من ٣٠ سنة	٣٠ سنة فأكثر
يوميًا	٢٢	٢٣	٣٠	٥٠
من ٥ إلى ٦ أيام	٣٠	٥٠	٣٠	٢٨
من ٣ إلى ٤ أيام	١٧	١٣	١٠	١٢
يوم أو يومان	٢١	٢٤	٢٥	١٥
المجموع	٩٠	١١٠	٩٥	١٠٥

فهذا الجدول يتضمن توزيع العينة حسب ثلاثة متغيرات (معدل قراءة الصحف، النوع، السن)، ويلاحظ أن مجموع تكرارات معدل قراءة الصحف ظل كما هو، كما أن مجموع الذكور، وكذلك مجموع الإناث ظل كما هو وإن كان قد توزع حسب السن. وأيّا كانت الجداول التكرارية، فإن هناك أنماطًا متعددة أخرى من الجداول يصممها الباحث حسب متطلبات النتائج سواء أكانت بسيطة أم معقدة، وقد لا يتضمن البحث جداول

تكرارية، لكنه يتضمن جداول من نوع آخر تلخص النتائج التي تم التوصل إليها، وعلى الرغم من ذلك تبقى الجداول التكرارية أساسية.

تصميم الجدول التكراري:

في معظم بحوث الإعلام والاتصال يتم الحصول على بيانات كمية (Quantitative Data) كثيرة ومتنوعة؛ لأنها مستمدة من عينات كبيرة، ومثل هذه البيانات لا يكون لها معنى دون تنظيمها في جداول تكرارية ذات معنى من خلال فئات واضحة، ولتوضيح ذلك نفرض أن لدينا بيانات عن المدة الزمنية (بالدقائق) والتي استغرقتها ٣٠ حلقة من برامج المقابلات التليفزيونية:

٧٨- ٦٧- ٣٠- ٥٤- ٣١- ٣٤- ٦٢- ٧٨- ٤٦ – ٧٠- ٦٠ – ٧٩ – ٧٠ – ٧٦- ٣٠ – ٤٥- ٦٠- ٣٨- ٥٦- ٥٦- ٧٨-
٤٧-٦٨- ٥٦- ٥٢- ٧٩-٦٥- ٦٣-٤٨-٤٧

من الواضح أن هذه البيانات غير ذات معنى، فهي لا توضح لنا عدد حلقات البرامج التي استغرقت مدة أطول أو أقل من زمن معين، كما أنها غير مرتبة تصاعديّا أو تنازليّا حتى يمكن معرفة شيء محدد، إنها مجرد كومة من الأرقام غير المفيدة. من هنا نحتاج إلى تبويبها في جدول واضح ومختصر حتى تكون ذات معنى ودلالة.

نقطة البدء في ذلك هي أن نحدد المدى (Range) الذي تشغله تلك القيم، وذلك من خلال تحديد أصغر قيمة وأكبر قيمة ، ثم نقسم هذا المدى على عدد الفئات الذي نراه مناسبًا. إن خارج هذه القسمة يعطينا أقرب قيمة صحيحة لطول الفئة. وبالنظر إلى البيانات المذكورة، نجد أن أصغر قيمة هي ٣٠ أما أكبر قيمة فهي ٨٠ ، وبالتالي يكون المدى هو ٨٠- ٣٠ = ٥٠ ، (وهناك من يضيف واحدًا صحيحًا إلى قيمة المدى).

إذا كان المدى هو ٥٠ فما هو طول الفئة المناسب؟ (أو ما هي سعة الفئة المناسبة؟) وما هو عدد الفئات المناسب؟ ليست هناك قاعدة ثابتة في هذا الشأن بالرغم من وجود بعض المعادلات التي تعتمد على لوغاريتم عدد المفردات مثل معادلة ستيرجس (Sturges)

والتي تتطلب الاستعانة بجداول اللوغاريتمات، وكذلك معادلة يول (Yole) التي تتطلب الاستعانة بجداول قوى وجذور الأعداد الطبيعية، لكن هذه المعادلات تقيد حرية الاختيار، وقد نحصل منها على عدد فئات غير مناسب، كأن تتركز معظم التكرارات في فئة معينة، ويكون هناك عدد محدود جدّا من التكرارات في فئات أخرى.

من هنا يتم قسمة المدى على عدد الفئات الذي نختاره، والذي نراه مناسبًا. في مثالنا هذا فإن المدى هو ٥٠، فإذا قرر الباحث أن يوزع البيانات المذكورة على خمس فئات، فإن طول الفئة يكون:

٥٠÷ ٥ = ١٠ أي أن سعة الفئة أو طول الفئة هو ١٠، وبناء على ذلك نبدأ الفئة الأولى بأصغر قيمة وهي ٣٠؛ ونظرًا لأن طول الفئة هو عشرة، فتكون الفئات على النحو الآتي:

٣٠ -٤٠ (من ٣٠ لأقل من ٤٠)

٤٠ - ٥٠ (من ٤٠ لأقل من ٥٠)

٥٠ - ٦٠ (من ٥٠ لأقل من ٦٠)

٦٠ - ٧٠ (من ٦٠ لأقل من ٧٠)

٧٠ - ٨٠ (من ٧٠ لأقل من ٨٠)

أي أن بداية الفئة الأولى - أو الحد الأدنى لها هو (٣٠) - وبذلك تستوعب أصغر تكرار، كما أن نهاية الفئة الأخيرة - أو الحد الأعلى لها- هو (٨٠)، وبذلك تستوعب أكبر تكرار، وكثيرًا ما يتم تحديد بداية الفئة الأولى بالرقم (صفر)، أو الرقم (٥)، ومن الشائع أيضًا تحديد الفئات بالرقم (١٠ ومضاعفاته) مع مراعاة أن تكون الفئات تستوعب جميع القيم (التكرارات)، فإذا كانت أصغر قيمة (٢٢) يمكن بدء الفئة الأولى من (٢٠)، وإذا كانت أكبر قيمة (٧٨) تكون نهاية الفئة الأخيرة (٨٠) ، ومع التأكيد على أن عملية التصنيف تتم باستخدام البرامج الإحصائية الجاهزة مثل (SPSS, SAS)، إلا أن الباحث

هو الذي يحدد عدد الفئات وطول الفئة، ومن المفترض أن يكون على معرفة بكيفية استخدام البرنامج في هذا التحديد.

وبموجب تحديد عدد الفئات وطول الفئة، يتم وضع كل رقم في الفئة المناسبة، فبالعودة إلى البيانات التي توضح المدة الزمنية لكل حلقة من حلقات البرامج التليفزيونية، نجد أن لدينا ثلاثين حلقة كانت مدة كل منها بالدقيقة هي:

٧٨- ٦٧- ٣٠- ٥٤- ٣١- ٣٤- ٦٢- ٧٨- ٤٦ – ٧٠- ٦٠ – ٧٩ – ٧٠- ٧٦- ٣٠ – ٤٥- ٦٠- ٣٨- ٥٦- ٥٦- ٧٨- ٤٧-٦٨- ٥٦- ٥٢- ٧٩-٦٥- ٦٣-٤٨-٤٧

أي أن الحلقة الأولى استغرقت ٧٨ دقيقة، وبالتالي فإنها تصنف في الفئة الأخيرة (من ٧٠ لأقل من ٨٠) أما الحلقة الثانية فقد استغرقت ٦٧ دقيقة، وبالتالي تصنف في الفئة الرابعة (من ٦٠ لأقل من ٧٠) وبخصوص الحلقة الثالثة فإنها استغرقت ٣٠ دقيقة، وبالتالي تصنف في الفئة الأولى (من ٣٠ لأقل من ٤٠)... وهكذا حتى تنتهي عملية تصنيف جميع الحلقات ضمن الفئة التي تنتمي إليها كل حلقة، وتكون كل حلقة بمثابة تكرار (Frequency)، على النحو المبين بالجدول الآتي:

الفئات	العلامات	التكرارات
٣٠ - ٣٩	///// (٣٠-٣١-٣٤-٣٠-٣٨)	٥
٤٠ - ٤٩	///// (٤٦-٤٥-٤٧-٤٨-٤٧)	٥
٥٠ -٥٩	///// (٥٤-٥٦-٥٦-٥٦-٥٢)	٥
٦٠ - ٦٩	///// // (٦٧-٦٢-٦٠-٦٠-٦٨-٦٥-٦٣)	٧
٧٠ إلى ٨٠	///// /// (٧٨-٧٨-٧٠-٧٩-٧٠-٧٦-٧٨-٧٩)	٨
المجموع	///// ///// ///// ///// // ///// ///	٣٠

ويمثل التكرار بشرطة مائلة (/) ولسهولة حساب أو عد هذه التكرارات يتم تمييزها في صورة مجموعات، كل مجموعة تضم خمسة تكرارات، وذلك إما بأن تكون العلامة الخامسة متقاطعة مع العلامات الأربع، أو بترك مسافة واضحة بين كل مجموعة تضم خمس علامات. هكذا يصبح لدينا جدول مختصر يوضح المعاني والدلالات التي تنطوي عليها البيانات السابقة، وذلك على الصورة الآتية:

الفئات	ك
٣٠ – ٤٠	٥
٤٠ -٥٠	٥
٥٠- ٦٠	٥
٦٠ – ٧٠	٧
٧٠ - ٨٠	٨
المجموع	**٣٠**

فإذا كانت فئات هذا الجدول تعبر عن المدة الزمنية، فإن التكرارات تعني عدد الحلقات موزعة حسب هذه الفئات، ويكشف الجدول عن أن هناك ١٥ حلقة؛ أي ما يعادل ٥٠% من حلقات البرامج كانت مدتها أقل من ساعة، أما النسبة الباقية (٥٠%) فقد كانت مدتها ساعة أو أكثر. ويمكن للباحث إعادة تنظيم الجدول بضم الفئات الثلاث الأولى في فئة واحدة؛ بحيث يتضمن الجدول ثلاث فئات فقط:

الفئات	ك
٣٠ – ٦٠	١٥
٦٠ – ٧٠	٧
٧٠ - ٨٠	٨
المجموع	**٣٠**

ومن الواضح أن الفئات قد اختلفت من حيث الطول، فالفئة الأولى طولها ثلاثون، أما الفئة الثانية فطولها (عشرة)، بينما طول الفئة الثالثة (أحد عشر)؛ أي أن فئات الجدول غير منتظمة من حيث الطول أو السعة.

الفئات المنتظمة والفئات غير المنتظمة:

فالفئات المنتظمة هي الفئات متساوية الطول. أما الفئات غير المنتظمة، فهي الفئات غير المتساوية في الطول، كمثال على الفئات المنتظمة نقدم الجدول الآتي، والذي يوضح توزيع عينة من قراء الصحف اليومية حسب عدد الأخبار التي أثارت انتباههم خلال عام كامل:

عدد الأخبار	التكرارات
٥-١٠	٩٨
١٠-15	٧٤
١٥-٢٠	٢٣
٢٠-٢٥	٥
المجموع	**٢٠٠**

يلاحظ من الجدول أنه يتضمن أربع فئات متساوية الطول، فكل فئة مداها خمسة، وأمام كل فئة عدد التكرارات الخاصة بها، فهناك ٩٨ شخصًا- أي ما يعادل ٤٩% من العينة- دلت استجاباتهم على أن الأخبار التي لفتت انتباههم خلال العام تتراوح ما بين خمسة لأقل من عشرة أخبار، كما أن ٢.٥% فقط - أي ما يعادل خمسة مفحوصين دلت استجاباتهم على أن الأخبار التي لفتت انتباههم تتراوح ما بين ٢٠ لأقل من ٢٥ خبرًا... وهكذا.

أما فيما يخص الفئات غير المنتظمة، فهي الفئات غير متساوية الطول، والجدول الآتي يوضح هذه الفكرة من واقع توزيع عينة من مستمعي برامج الراديو حسب أعمارهم بالسنوات:

التكرارات	فئات الأعمار
٢٤٥	٢٠- ٢٩
٨٤	٣٠- ٤٩
٢١	٥٠- ٥٤
٣٥٠	**المجموع**

من الواضح أن فئات الأعمار غير متساوية الطول، فالفئة الأولى طولها ١٠ (إذ إنها من ٢٠ وتنتهي عند ٢٩)، أما الفئة الثانية فطولها ٢٠ (إذ إنها تمتد من ٣٠ إلى ٤٩)، والفئة الثالثة طولها ٥ (إذ إنها تمتد من ٥٠ إلى ٥٤).

الفئات المغلقة والفئات المفتوحة:

الفئات المغلقة (Closed Categories)، هي الفئات التي لها بداية ونهاية واضحة، محددة، ففي الجدول السابق نجد أن جميع الفئات مغلقة، فالفئة الأولى تبدأ من عشرين وتنتهي بتسعة وعشرين، أما الفئة الثانية فتبدأ من ثلاثين وتنتهي بتسعة وأربعين، بينما الفئة الثالثة تبدأ من خمسين وتنتهي بأربعة وخمسين.

مقابل ذلك هناك الفئات المفتوحة (Open Categories)، وهي قد تكون مغلقة البداية ومفتوحة النهاية أو العكس: مفتوحة البداية ومغلقة النهاية. كمثال للجداول ذات الفئات مغلقة البداية ومفتوحة النهاية، هذا الجدول الذي يتضمن توزيع عينة قوامها ١٠٠٠ مفردة من جمهور الراديو حسب الأعمار:

التكرارات	فئات الأعمار
٤٦٩	٢٠-
٣٠٣	٤٠-
١٨٤	٥٠-
٤٤	٦٠-
١٠٠٠	**المجموع**

واضح من هذا الجدول أن فئات الأعمار مفتوحة النهايات، وباستثناء الفئة الأخيرة، فإن نهايات بقية الفئات مفهومة ضمنًا، فالفئة الأولى تبدأ من عشرين سنة وإن كانت نهايتها غير محددة، غير أن نهايتها معروفة من واقع بداية الفئة اللاحقة والتي تبدأ من أربعين، فكأن نهاية الفئة الأولى هي ٣٩. المنطق نفسه فيما يخص الفئة الثانية (نهايتها ٤٩)، والفئة الثالثة نهايتها ٥٩ غير أن الفئة الأخيرة – والتي تبدأ من ٦٠ - فإن نهايتها مفتوحة، فهي لا يعقبها فئة ذات بداية محددة، وهناك بعض المقاييس الإحصائية- حسبما سنوضح في موضع لاحق- لا يمكن استخدامها مع التكرارات التي تتوزع على فئات مفتوحة.

النوع الآخر من الفئات المفتوحة يتمثل في الفئات مفتوحة البداية ومغلقة النهاية، يتضح ذلك في الجدول الآتي والذي يتضمن توزيع عينة من مشاهدي التليفزيون (٣٥٠) مفردة حسب درجاتهم على مقياس للمعلومات العامة (الدرجة الكلية للمقياس ٣٠ درجة):

فئات الدرجات	التكرارات
١٥-	٥١
-١٥	٢٠٢
-٢٠	٥٦
٣٠-٢٥	٤١
المجموع	**٣٥٠**

يعرف هذا النمط بأنه الجدول المفتوح من أعلى، فالفئة الأولى تعني أقل من ١٥ (فهي غير محددة البداية)، فالمفحوصون الذين حصلوا على أقل من ١٥ درجة، يقعون على المتصل من صفر حتى ١٤.٥ درجة، بمعنى أن بعضهم قد حصل على صفر، والبعض الآخر حصل على ٢ والبعض حصل على ١٠ وهكذا، أما بقية الفئات، فإن نهاياتها مفهومة، فالفئة الثانية نهايتها أقل من ٢٠، والفئة الثالثة نهايتها أقل من ٢٥، أما الفئة الأخيرة، فإن نهايتها (٣٠).

وهناك جداول تكرارية مفتوحة البداية (بداية الفئة الأولى)، ومفتوحة النهاية (نهاية الفئة الأخيرة)، وبالمثل، هناك جداول تكرارية مغلقة البداية (بداية الفئة الأولى)، ومغلقة النهاية (نهاية الفئة الأخيرة)، ومن الأفضل في تصميم الجدول التكراري أن تكون الفئات مغلقة البداية ومغلقة النهاية، وبوجه عام، فإن تصميم الجدول التكراري الأمثل هو الذي تتوافر فيه الخصائص الآتية:

- أن تكون الفئات مستقلة، غير متداخلة، بمعنى أن تبدأ الفئة اللاحقة من حيث انتهت السابقة عليها.
- أن يكون عدد الفئات معقولًا، فلا يتضمن الجدول فئات كثيرة لكل منها تكرارات محدودة، ولا يتضمن فئات قليلة جدّا تتركز فيها معظم التكرارات.
- تحاشي استخدام الفئات المفتوحة، وكذلك تحاشي الفئات غير المنتظمة قدر الإمكان.
- وضع عنوان واضح ومختصر للجدول، بما يعكس محتواه بدقة.

التوزيع التكراري المتجمع:

نستخدم التوزيع التكراري المتجمع بهدف معرفة عدد التكرارات التي تقل أو التي تزيد عن قيمة معينة. وهناك التوزيع التكراري المتجمع الصاعد، كما أن هناك التوزيع التكراري المتجمع الهابط

(أ) التوزيع التكراري المتجمع الصاعد:

عندما تتوزع التكرارات حسب فئات (Categories) كميّة، لكل فئة سعة معينة أو مدى معين، فإن ذلك يعني أن كل فئة من فئات الجدول التكراري لها حد أدنى وحد أعلى، والمنحنى التكراري المتجمع الصاعد هو عدد التكرارات التي تقع دون الحد الأعلى للفئات. لتوضيح ذلك نفرض أن لدينا (٥٠) نشرة أخبار، كانت مدة كل نشرة بالدقيقة على النحو التالي:

٤١	٣٣	٢٠	٢٣	٤١
٣٥	٢٥	٢٣	٢٤	٣٩
٣١	٢٨	٣١	٢٧	٢٠
٤٣	٢٩	٣٠	٤١	٣٤
٤٨	٣٥	٣٧	٣٢	٣٥
٤٧	٢٨	٤٢	٢٨	٢٩
٤١	٣٣	٢٥	٤١	٢١
٣٨	٣٢	٢٥	٣٦	٢٦
٣٥	٣٩	٢٤	٣٧	٢٨
٢٩	٤١	٣٢	٤٩	٣٩

فالنشرة الأولى مدتها (٤١) دقيقة، وإذا اتجهنا من الخلية الأولى رأسيّا إلى أسفل، فإن النشرة التالية لها مدتها (٣٥) دقيقة، أما النشرة الأخيرة فإن مدتها (٣٩) دقيقة.. وهكذا. وللحصول على التوزيع التكراري المتجمع الصاعد يتعين أولا الحصول على التوزيع التكراري البسيط، ونبدأ بتحديد عدد الفئات بحيث تستوعب جميع القيم (المدد الزمنية)، ومن الواضح أن أقل نشرة، كانت مدتها (٢٠) دقيقة، أما أطول نشرة فكانت مدتها (٤٩) دقيقة، ويمكن للباحث تحديد عدد الفئات الذي يراه مناسبًا، ويجب أن تكون بداية الفئة الأولى تساوي أصغر قيمة (أو تقل عنها بواحد صحيح)، وأن تكون نهاية الفئة الأخيرة تساوي أكبر قيمة (أو تزيد عنها بواحد صحيح)، فإذا حددنا عدد الفئات كأن يكون ست فئات مثلًا، ووضعنا التكرارات المقابلة لهذه الفئات، فإننا نحصل على الجدول التكراري التالي:

فئات الوقت بالدقيقة	ك
٢٠ لأقل من ٢٥	٧
٢٥ لأقل من ٣٠	١٢
٣٠ لأقل من ٣٥	٩
٣٥ لأقل من ٤٠	١١
٤٠ لأقل من ٤٥	٨
٤٥ لأقل من ٥٠	٣
المجموع	**٥٠**

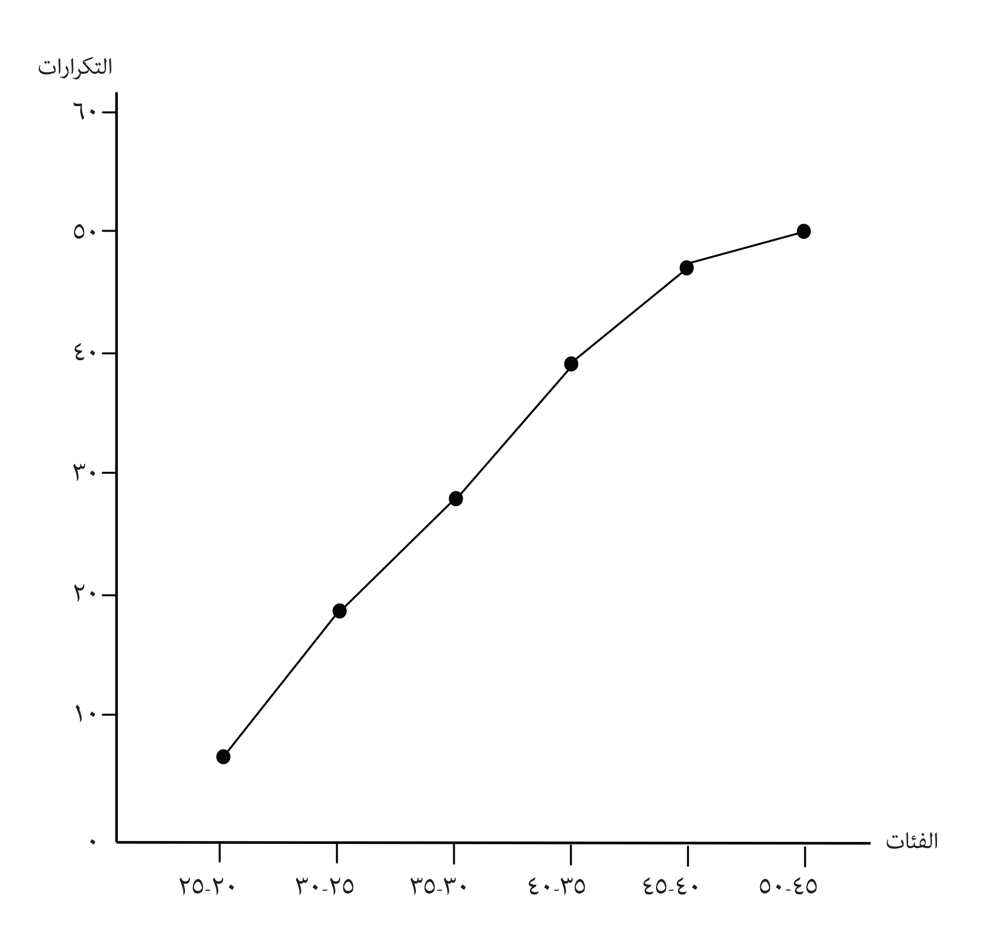
هذا هو الجدول التكراري البسيط، ومنه نتبين أن الفئة الأولى بها سبعة تكرارات؛ أي أن هناك سبع نشرات تقل مدة كل منها عن (٢٥) دقيقة، أما الفئة الثانية، فإن بها (١٢) تكرارًا، أي أن هناك (١٢) نشرة إخبارية تتراوح مدة كل منها بين ٢٥ دقيقة لأقل من ٣٠ دقيقة .... وهكذا، ومن التوزيع التكراري البسيط يمكن بسهولة الحصول على التوزيع التكراري المتجمع الصاعد من خلال رصد التكرارات التي تقل عن الحد الأعلى لكل فئة، فالفئة الأولى مثلًا هي (من ٢٠ لأقل من ٢٥)؛ أي أن حدها الأعلى هو أقل من (٢٥)، وبالتالي، فإن التوزيع التكراري المتجمع الصاعد لهذه الفئة سيشمل جميع التكرارات (عدد النشرات) التي تقل مدة كل منها عن ٢٥ دقيقة، والفئة الثانية هي (من ٢٥ لأقل من ٣٠)، أي أن حدهاالأعلى هو أقل من (٣٠) دقيقة، وبالتالي فإن التوزيع التكراري المتجمع الصاعد لهذه الفئة سيشمل جميع التكرارات التي تقل عن (٣٠)؛ أي أن التكرارات المتجمعة للفئة الثانية تشمل تكرارات الفئة الأولى مضافًا إليها تكرارات الفئة الثانية.. وهكذا فيما يخص جميع الفئات، فنحصل على الجدول التالي:

فئات الوقت بالدقيقة	ك	الحدود الأعلى للفئات	التوزيع التكراري المتجمع الصاعد
٢٠ لأقل من ٢٥	٧	٢٥-	٧
٢٥ لأقل من ٣٠	١٢	٣٠-	١٩
٣٠ لأقل من ٣٥	٩	٣٥-	٢٨
٣٥ لأقل من ٤٠	١١	٤٠-	٣٩
٤٠ لأقل من ٤٥	٨	٤٥-	٤٧
٤٥ لأقل من ٥٠	٣	٥٠-	٥٠
المجموع	**٥٠**		

لقد تم تكوين المنحنى التكراري المتجمع الصاعد (ت. م. ص.) من خلال الحدود الأعلى للفئات (العمود الثالث)، والتكرارات المناظرة لها. ومن الجدول نتبين أن:

- التوزيع التكراري المتجمع الصاعد للفئة الأولى يساوي عدد تكرارات هذه الفئة (٧) تكرارات.
- التوزيع التكراري المتجمع الصاعد للفئة الثانية يساوي عدد تكرارات الفئة الأولى مضافًا إليه عدد تكرارات الفئة الثانية (٧+ ١٢= ١٩).
- التوزيع التكراري المتجمع الصاعد للفئة الثالثة يساوي عدد تكرارات الفئة الأولى + تكرارات الفئة الثانية + تكرارات الفئة الثالثة (٧+ ١٢ + ٩= ٢٨).
- التوزيع التكراري المتجمع الصاعد للفئة الرابعة يساوي عدد تكرارات الفئات الأولى والثانية والثالثة والرابعة (٧+ ١٢ + ٩+ ١١= ٣٩).
- التوزيع التكراري المتجمع الصاعد للفئة الخامسة يساوي عدد تكرارات الفئات الأولى والثانية والثالثة والرابعة والخامسة (٧+ ١٢ + ٩+ ١١ +٨= ٤٧).
- التوزيع التكراري المتجمع الصاعد للفئة السادسة يساوي عدد تكرارات هذه الفئة مضافًا إليها تكرارات كل الفئات السابقة (٧+ ١٢ + ٩+ ١١ +٨ + ٣= ٥٠).

وينتهي التوزيع التكراري المتجمع الصاعد بمجمل التكرارات (ففي المثال المبين نلاحظ أن جميع نشرات الأخبار محل الدراسة تقل مدتها عن ٥٠ دقيقة، وهذه القيمة هي الحد الأعلى للفئة الأخيرة). ويمكن رسم المنحنى المتجمع الصاعد (م. ت. ص.) باستخدام الفئات والتكرارات الخاصة بها؛ حيث يتم تمثيل الفئات بمحور أفقي بينما يتم تمثيل التكرارات بمحور رأسي، ويتم وضع نقطة تقابل التكرارات الخاصة بكل فئة؛ ومن ثمّ يتم توصيل هذه النقاط فيكون لدينا المنحنى المتجمع الصاعد على النحو التالي:



ومن التوزيع التكراري المتجمع الصاعد، يمكن الحصول على التوزيع المتجمع الصاعد المئوي أو النسبي، وذلك من خلال حساب النسبة المئوية لتكرارات كل فئة من مجمل التكرارات، فالتكرار المئوي للفئة الأولى = ٧÷ ٥٠× ١٠٠ = ١٤%.... وهكذا، كما يتضح ذلك في العمود الأخير من الجدول الآتي:

فئات الوقت بالدقيقة	ك	التوزيع التكراري المتجمع الصاعد	التوزيع المتجمع الصاعد المئوي
٢٠ لأقل من ٢٥	٧	٧	١٤
٢٥ لأقل من ٣٠	١٢	١٩	٣٨
٣٠ لأقل من ٣٥	٩	٢٨	٥٦
٣٥ لأقل من ٤٠	١١	٣٩	٧٨
٤٠ لأقل من ٤٥	٨	٤٧	٩٤
٤٥ لأقل من ٥٠	٣	٥٠	١٠٠
المجموع	**٥٠**		

ومن خلال هذا التوزيع نتبين أن ٩٤% من نشرات الأخبار عينة البحث تقل مدة كل منها عن ٤٥ دقيقة، كما أن ٣٨% من تلك النشرات تقل مدة كل منها عن ٣٠ دقيقة، كما أن جميع تلك النشرات تقل مدة كل منها عن ٥٠ دقيقة..... وهكذا.

(ب) التوزيع التكراري المتجمع الهابط:

وهو عكس التوزيع التكراري المتجمع الصاعد، ذلك أن التوزيع التكراري المتجمع الهابط (ت. م. هـ) تبدأ تكراراته بالتكرارات التي تزيد على الحد الأدنى للفئات، وبالتالي فإنه يبدأ بإجمالي التكرارات، ففي المثال السابق نجد أن الحد الأدنى للفئة الأولى هو (٢٠)، ومن الواضح أن جميع التكرارات (وهي خمسون تكرارًا) تزيد عن (٢٠) دقيقة، بمعنى أن جميع التكرارات توضع ضمن أول فئة للتوزيع التكراري المتجمع الهابط (لأن جميع النشرات الإخبارية تزيد مدة كل منها على عشرين دقيقة)، أما الفئة الثانية فحدها الأدنى (٢٥) دقيقة، وفي الوقت نفسه نجد أن التكرارات الأصلية للفئة الأولى (عدد النشرات) هو (٧)؛ أي أن هناك سبع نشرات - هي تكرارات الفئة الأولى- تقل مدة كل منها عن (٢٥) دقيقة، وبالتالي يكون التوزيع التكراري المتجمع الهابط والخاص بالفئة الثانية =

٥٠- ٧ = ٤٣.... وهكذا، فإن التكرارات يقل عـددها في الفئـات الأكـبر تـدريجيّا إلى أن تـصل إلى الحـد الأدنى أو الصفر في الفئة الأخيرة. فإذا أردنا تكوين التوزيع التكراري المتجمع الهـابط (ت.م. هــ) ككـل، فإننا نبدأ بالفـئة الأولى ويقع في حدها الأدنى جميع التكرارات، أما الفئـة الثانيـة فتـشمل هـذا المجمـل منقوصًا منه تكرارات الفئة الأولى؛ أي أننا نطرح (ننقص) تكرارات كل فئة من تكـرارات الفئـة الـسابقة عليها على النحو المبين بهذا الجدول:

فئات الوقت بالدقيقة	ك	الحدود الأدنى للفئات	التوزيع التكراري المتجمع الهابط
٢٠ لأقل من ٢٥	٧	٢٠-	٥٠
٢٥ لأقل من ٣٠	١٢	٢٥-	٤٣
٣٠ لأقل من ٣٥	٩	٣٠-	٣١
٣٥ لأقل من ٤٠	١١	٣٥-	٢٢
٤٠ لأقل من ٤٥	٨	٤٠-	١١
٤٥ لأقل من ٥٠	٣	٤٥-	٣
المجموع	٥٠	٥٠-	صفر

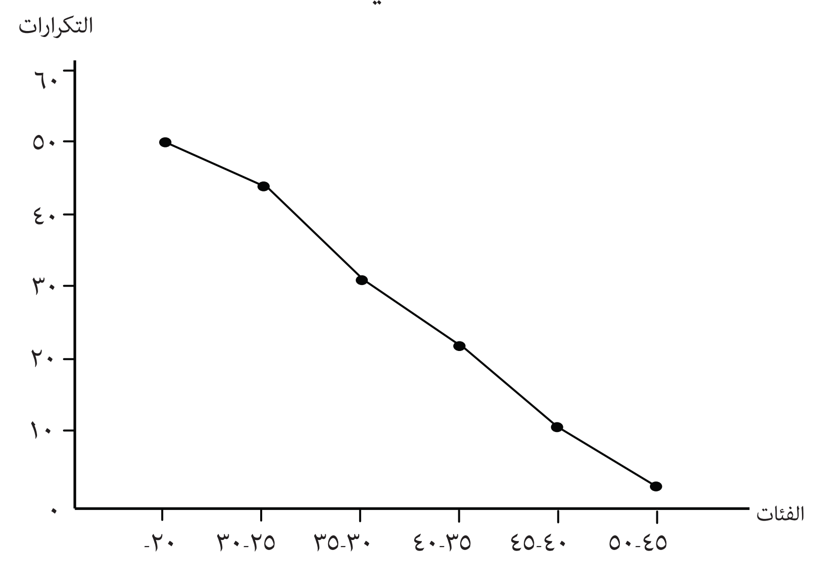
لقد تم تكوين المنحنى التكراري المتجمع الهابط (ت. م. هــ) مـن خـلال الحـدود الـدنيا للفئـات (العمود الثالث)، فالحد الأدنى للفئة الأولى هو (٢٠)، وبالتالي فإن كل التكرارات التي تزيد على عـشرين هي التكرار المتجمع الهابط والخاص بالفئة الأولى، كما أن الحد الأدنى للفئة الثانيـة هـو (٢٥)، وبالتـالي فإن كل التكرارات التي تزيد على خمـسة وعـشـرين هـي المنحنـى التكـراري المتجمـع الهـابط الخـاص بالفئة الثانية.... وهكذا، تكون التكرارات المبينة في العمود الأخير هي التوزيع التكراري المتجمع الهابط.
ومن الجدول نتبين أن:

- التوزيع التكراري المتجمع الهابط للفئة الأولى يساوي إجمالي التكرارات (إذ إن جميع نشرات الأخبار محل الدراسة تبلغ مدة كل منها عشرين دقيقة أو أكثر).
- التوزيع التكراري المتجمع الهابط للفئة الثانية يساوي مجمل التكرارات منقوصًا منه تكرارات الفئة الأولى (٥٠- ٧ = ٤٣).
- التوزيع التكراري المتجمع الهابط للفئة الثالثة يساوي مجمل التكرارات منقوصًا منه تكرارات الفئة الأولى وتكرارات الفئة الثانية (٥٠-٧- ١٢= ٣١).
- التوزيع التكراري المتجمع الهابط للفئة الرابعة يساوي مجمل التكرارات منقوصًا منه تكرارات الفئات الأولى والثانية والثالثة (٥٠-٧- ١٢- ٩= ٢٢).
- التوزيع التكراري المتجمع الهابط للفئة الخامسة يساوي مجمل التكرارات منقوصًا منه تكرارات الفئات من الأولى حتى الرابعة (٥٠-٧- ١٢- ٩ – ١١= ١١).
- التوزيع التكراري المتجمع الهابط للفئة السادسة يساوي مجمل التكرارات منقوصًا منه تكرارات الفئات من الأولى حتى الخامسة (٥٠-٧- ١٢- ٩ – ١١-٨= ٣).

وينتهي التوزيع التكراري المتجمع الهابط بأصغر قيمة أو بقيمة الصفر؛ حيث تكون الفئات قد استوعبت مجمل التكرارات (٥٠- ٥٠ = صفر)، ففي المثال المبين لا توجد نشرات أخبار مدتها خمسون دقيقة فأكثر.

وباستخدام التكرارات يمكن رسم المنحنى المتجمع الهابط (م. ت. هـ) بيانيًا، حيث نرسم خطًا أفقيًا لتمثيل الفئات، وخطًا آخر رأسيًا لتمثيل التكرارات المقابلة لكل فئة، فبالنسبة للفئة الأولى مثلا (٢٠ – ٢٥) والمبينة على المحور الأفقي نتجه إلى أعلى ونضع نقطة مقابلة لتكرارات هذه الفئة (٥٠ تكرارًا)، وهكذا فيما يخص بقية الفئات، وبتوصيل هذه النقاط ينتج لنا المنحنى التكراري المتجمع الهابط (م. ت. هـ) على النحو التالي:





ويمكن الحصول على التوزيع التكراري المتجمع الهابط المئوي أو النسبي، وذلك من خلال حساب النسبة المئوية لتكرارات كل فئة من مجمل التكرارات، فالتكرار المئوي للفئة الأولى = ٥٠÷ ٥٠× ١٠٠ = ١٠٠%، أما التكرار المئوي للفئة الثانية فهو ٤٣÷ ٥٠ = ٨٦% وهكذا، كما يتضح ذلك في العمود الأخير من الجدول الآتي:

فئات الوقت بالدقيقة	ك	التوزيع لتكراري المتجمع الهابط	التوزيع المتجمع الهابط المئوي
٢٠ لأقل من ٢٥	٧	٥٠	١٠٠
٢٥ لأقل من ٣٠	١٢	٤٣	٨٦
٣٠ لأقل من ٣٥	٩	٣١	٦٢
٣٥ لأقل من ٤٠	١١	٢٢	٤٤
٤٠ لأقل من ٤٥	٨	١١	٢٢
٤٥ لأقل من ٥٠	٣	٣	٦
المجموع	**٥٠**	**صفر**	**صفر**

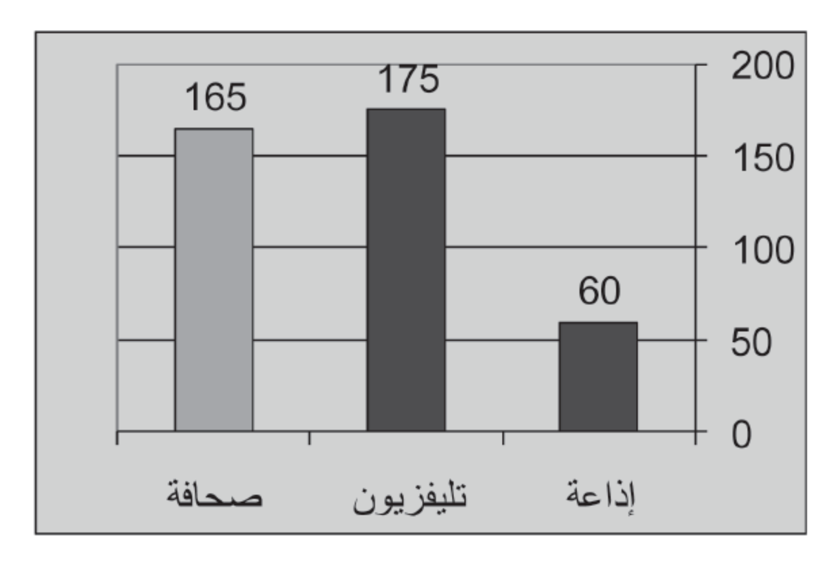
فمن خلال هذا التوزيع نتبين مثلًا أن ٦٢% من نشرات الأخبار مجال البحث تقل مدة كل منها عن ٣٥ دقيقة، كما أن ٢٢% من تلك النشرات تقل مدة كل منها عن ٤٥ دقيقة.. وهكذا.

ثانيًا: الرسوم البيانية

تتعدد الرسوم البيانية التي يمكن استخدامها في التعبير عن المعطيات الإحصائية، ومن أبرزها: الأعمدة، المدرج التكراري، المضلع التكراري، المنحنيات، الدوائر... إلخ، وكل شكل من هذه الأشكال يتخذ أنماطًا متعددة، ويمكن تطويعه بأساليب متنوعة لتمثيل المعطيات الإحصائية. وفي الصفحات القادمة تعريف موجز بأهم أساليب العرض البياني مع ربط ذلك ببحوث الإعلام والاتصال بالجماهير:

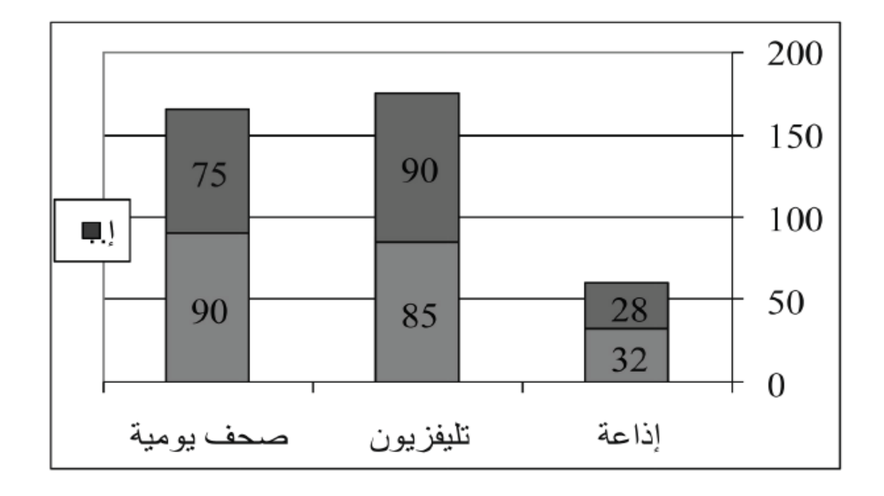
(أ) الأعمدة:

الأعمدة (Bar Charts) هي خطوط رأسية مستقيمة يتم رسمها على خط أفقي؛ بحيث تعكس تكرارات الظاهرة أو الموضوع. كمثال توضيحي نفرض أن لدينا عينة قوامها ٤٠٠ مفردة موزعة حسب مجال العمل (Work field) (إذاعة، تليفزيون، صحف يومية)؛ حيث يعمل (٦٠) في الإذاعة، (١٧٥) في التليفزيون، مقابل (١٦٥) في الصحف اليومية. للتعبير عن هذه النتيجة بالأعمدة نقوم برسم خط أفقي (Horizontal) يوضح مجالات العمل (إذاعة، تليفزيون، صحافة)، ومن طرف هذا الخط يتم رسم خط رأسي(Vertical) بمقياس رسم مناسب (يستوعب جميع الأرقام التي تعبر عن أعداد العاملين)، وعند النقطة التي تمثل الإذاعة نضع علامة ونتجه منها بخط مستقيم إلى أعلى بموازاة المحور الرأسي حتى نصل إلى النقطة التي تقابل القيمة (٦٠) وهو عدد العاملين بالإذاعة، ونكرر العملية نفسها فيما يخص أعداد العاملين بكل من الصحافة والتليفزيون، فيكون لدينا الرسم البياني الآتي:



فالأعمدة تعكس توزيع العينة حسب مجال العامل بدقة شديدة ووضوح؛ إذ إن مجرد النظر إلى تلك الأعمدة، يعطي صورة كاملة عن هذا التوزيع.

من جهة أخرى يمكن استخدام الأعمدة في تمثيل توزيع العينة حسب متغيرين: متغير مجال العمل (إذاعة، تليفزيون، صحافة)، ومتغير النوع (ذكور & إناث)، وتسمى في هذه الحالة بالأعمدة المزدوجة، وذلك على النحو الآتي:



فكل عمود يتضمن قسمين، القسم الأعلى يوضح عدد الإناث (Female)، والقسم الأسفل يوضح عدد الذكور (Male) في كل مجال، ومن الواضح تقارب أعداد الجنسين على مستوى كـل مجـال مـن المجـالات الثلاثة، وعلى يسار الأعمدة يوجد المفتاح الذي يدل على كل من الذكور والإناث.

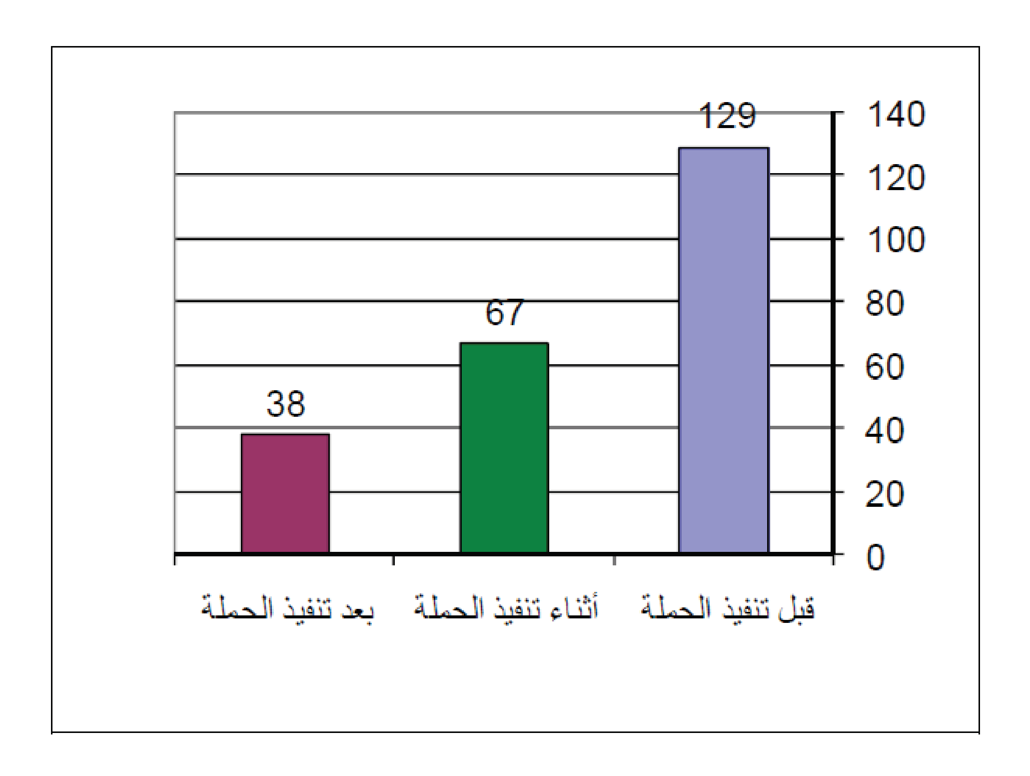
كما تستخدم الأعمدة للمقارنة بين الظاهرة عبر فترات زمنية ذات دلالة، مثال ذلك أن إحدى الحملات الإعلامية استهدفت تعريف النساء في المجتمع المحلي بحق المرأة في التعليم باعتباره أحد الحقوق التي كفلها الإسلام للمرأة، وقد تم رصد نسبة النساء اللاتي لا يعرفن معنى «حق المرأة في التعليم» قبل وأثناء وبعد تنفيذ حملة إعلامية من خلال التليفزيون تستهدف التعريف بهذا الحق، وكانت العينة (٢٣٤) مفردة، وقد تبين أن أعداد النساء اللاتي (لا يعرفن) هذا الحق كانت كالآتي:

١٢٩ (قبل تنفيذ الحملة)

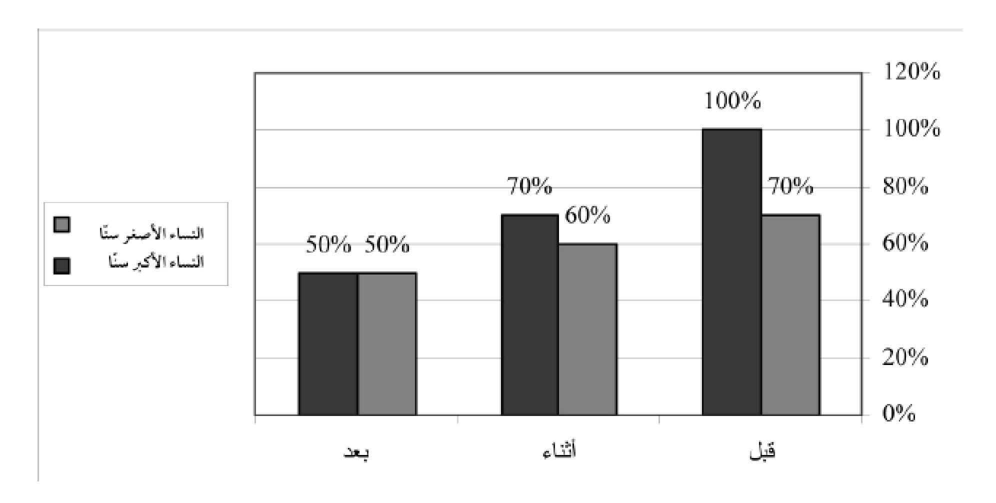
٦٧ (أثناء تنفيذ الحملة)

٣٨ (بعد تنفيذ الحملة)

وللتعبير عن هذه النتيجة بالأعمدة نقوم برسم خط رأسي (Vertical) بمقياس رسم مناسب (يستوعب جميع التكرارات التي تعبر عن عدد النساء اللاتي لا يعرفن معنى حق المرأة في التعليم)، وخط آخر أفقي (Horizontal) يوضح مراحل تنفيذ الحملة (قبل، أثناء، بعد)، وعند النقطة التي توضح مرحلة قبل الحملة - نضع علامة على المحور الأفقي ونرسم منها خطًا مستقيمًا يتجه إلى أعلى بموازاة المحور الرأسي حتى نصل إلى النقطة التي تقابل القيمة ١٢٩ على المحور الرأسي، وهذه النقطة تمثل عدد اللاتي (لا يعرفن) حق المرأة في التعليم قبل الحملة، ونكرر العملية نفسها فيما يخص أعداد النساء اللاتي (لا يعرفن) حق المرأة في التعليم أثناء تنفيذ الحملة، وبعد تنفيذ الحملة، فنحصل على الرسم البياني التالي:



وبالطريقة نفسها نستخدم الأعمدة لتمثيل أكثر من مجموعة يشملها القطاع الواحد. كمثال توضيحي، فإن إحدى الحملات الإعلامية استهدفت تعريف النساء بحقهن في «المشاركة السياسية»، وتم رصد نسبة النساء اللاتي (لا يعرفن) هذا الحق قبل وأثناء و بعد تنفيذ حملة إعلامية من خلال التليفزيون للتعريف بهذا الحق. فإذا أردنا تمثيل هذه النتيجة بيانيّا مع التمييز بين النساء الأصغر سنّا (مجموعة أولى) والنساء الأكبر سنّا (مجموعة ثانية)، فإننا نقوم بالإجراءات السابقة، ولكن مع تخصيص عمود يميز كل مجموعة، فنحصل على الرسم البياني التالي:



إن مجرد النظر إلى تلك الأعمدة يكشف عن أنه - قبل تنفيذ الحملة – كان ١٠٠% من النساء الأكبر سنّا لا يعرفن معنى «حق المشاركة السياسية»، انخفضت تلك النسبة إلى ٧٠% أثناء تقديم برامج ومواد الحملة، ومع نهاية الحملة، أصبحت تلك النسبة ٥٠ % تقريبا. هذا يعني أنه مع نهاية الحملة الإعلامية انخفض عدد النساء الأكبر سنّا اللاتي يجهلن حق المرأة في المشاركة السياسية (وهذا يعني زيادة عدد النساء اللاتي أصبحن يعرفن هذا الحق).

أما فيما يخص النساء الأصغر سنّا، فإن مجرد النظر إلى الشكل السابق يوضح أنه قبل تنفيذ الحملة كان ٧٠% من النساء لا يعرفن حقهن في المشاركة السياسية، انخفضت هذه النسبة إلى ٦٠% أثناء تنفيذ الحملة، ثم إلى ٥٠% بعد انتهاء الحملة. أي أن الحملة التي قدمها التليفزيون، قد ساهمت في تخفيض نسبة النساء اللاتي يجهلن حقهن في المشاركة السياسية من ١٠٠% إلى ٥٠% من مجموعة النساء الأكبر سنّا، ومن ٧٠% إلى ٥٠% من مجموعة النساء الأصغر سنّا، ويوضح الجزء الأخير من الرسم تساوي نسبة عدم المعرفة بين المجموعتين، بمعنى أنه بعد انتهاء الحملة، تبين أن هناك ٥٠% من المجموعتين، لا يعرفن أن النساء لهن الحق في المشاركة السياسية، وهنا يمكن استنتاج أن الحملة كانت

فعالة (بدرجة متوسطة)، لكنها حققت استفادة أكبر للنسـاء الأكبر سنّا مقارنة بالنساء الأصغر سنّا، (فقد انخفضت نسبة عدم المعـرفة من ١٠٠% إلى ٥٠% في المجموعة الأولى بينما انخفضت تلك النسبة من ٧٠% إلى ٥٠% في المجموعة الثانية).

(ب) المدرج التكراري:

المدرج التكراري(Histogram) عبارة عن عدد من المستطيلات كل منها بارتفاع معين (الطول) ليعـبر عن تكرارات الفئة، وباتساع معين (العرض) ليعبر عـن مـدى أو سـعة الفئـة، وبالتـالي يكـون عـدد مـن المستطيلات يساوي عدد الفئات (مستطيل لكل فئة). ولتوضـيح ذلـك نفـترض أن لـدينا الجـدول التـالي الذي يوضح توزيع عدد من المقـالات الـصحفية المنـشورة حـول قـضية معينـة حـسب فئـات المـساحة بالسنتيمتر المربع:

فئات المساحة (سم^2)	**عدد المقالات**
٥٠- ٦٠	٣٨
٦٠- ٧٠	٢٨
٧٠ إلى٨٠	١٤
المجموع	**٨٠**

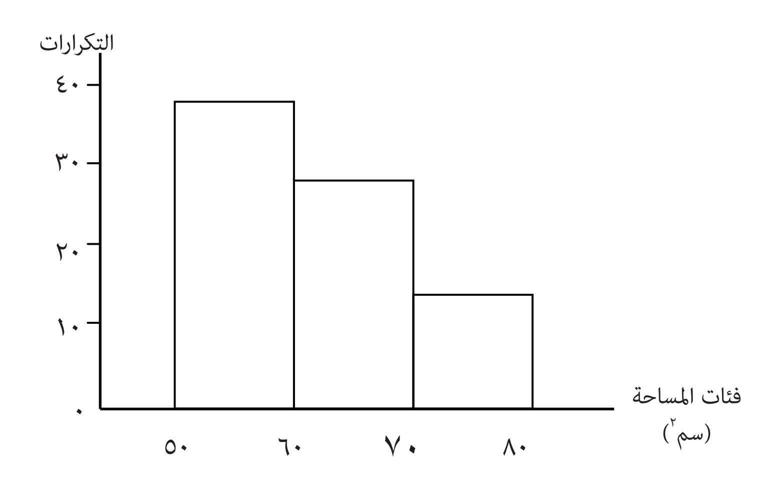
فهناك (٣٨) مقالًا صحفيّا تتراوح مساحة كل منها ما بين ٥٠ لأقل من ٦٠سـم^2 ، وهنـاك ٢٨ مقـالًا صحفيّا تتراوح مساحة كل منها ما بين ٦٠ لأقل من ٧٠سـم^2، كـما أن هنـاك ١٤ مقـالًا صـحفيّا تـتراوح مساحة كل منها ما بين ٧٠ إلى ٨٠ سم^2.

ويتم رسم مدرج تكراري يعبر عن هذا التوزيع من خلال رسم خط أفقي يمثل فئـات المساحة بالسنتيمتر المربع، وخط آخر رأسي يمثـل عـدد المقـالات الـصحفية الخاصـة بكـل فئة، أي أننا نحدد حدود الفئات على المحور الأفقي، ثم نحدد عدد التكـرارات - بمقيـاس رسـم مناسـب – عـلى المحـور الـرأسي. إن مقيـاس الرسـم المناسـب هـو الـذي يـستوعب

التكرارات بحديها الأدنى والأعلى، وفي مثالنا المذكور فإن الحد الأدنى هو (١٤) أما الحدالأعلى فهـو (٣٨) وبناء على ذلك فإن الخط الرأسي يبدأ بالصفر مرورًا بأقرب رقم عشري للقيمة (١٤)، وينتهي هذا الخط بأقرب رقم عشـري للقيمة (٣٨)؛ أي أن الخط الرأسي يبدأ بالصفر مرورًا بالرقم (١٠)، وينتهـي بالقيمـة (٤٠)، وعلى الخط الأفقي نقوم بتحديد الفئات المبينة بالجدول وهي:

- ٥٠ – ٦٠
- ٦٠- ٧٠
- ٧٠- ٨٠

ومن منتصف كل فئة من الفئات نتجه إلى أعلى ونضع نقطة تقابل تكرارات كل فئة على المحـور الرأسي ثم نرسم المستطيل من بداية الحد الأدنى للفئة وحتى الحدالأعلى لها، فنحصل على مجموعة من المستطيلات الرأسية، كل منها يعبر عن فئة بتكراراتها على النحو الذي يوضحه الشكل التالي:



فالمسـتطيل الأول يرتفع طوله إلى ٣٨ أما عرضه فيشمل الفئة الأولى (٥٠- ٦٠)، ويمثل المستطيل الثاني تكـرارات الفئـة الثانيـة؛ حيـث يرتفـع طولـه إلى ٢٨ أمـا عرضـه فيـشمل

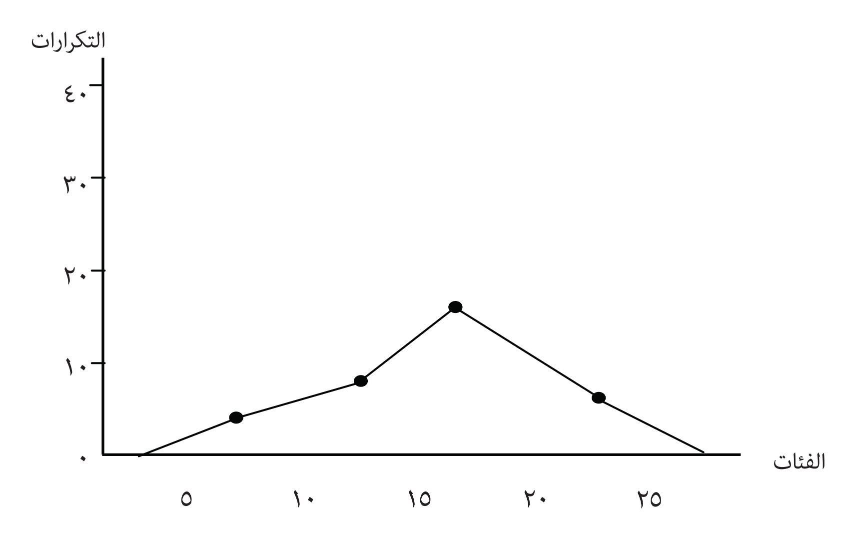
الفئة الثانية (٦٠-٧٠)، وأخيرًا فإن المستطيل الثالث يمثل تكرارات الفئة الثالثة؛ حيث يرتفع طوله إلى ١٤ أما عرضه فيشمل الفئة الثالثة (٧٠- ٨٠)، وهذا هو المدرج التكراري للجدول السابق (مع ملاحظة أن ارتفاع المستطيلات يتحدد بعدد تكرارات الفئة التي يمثلها).

(جـ) المضلع التكراري:

المضلع التكراري (Polygon) عبارة عن خط (Line) يصل بين التكرارات حسب مراكز الفئات، هذا الخط يغلق التوزيع من الناحيتين اليمنى واليسرى. ولتوضيح ذلك نفترض أن لدينا الجدول التالي الذي يوضح توزيع عدد من البرامج السياحية موزعة حسب المدة الزمنية بالدقيقة كما قدمها التليفزيون خلال دورة برامجية كاملة:

فئات المدة الزمنية	عدد البرامج
٥- ١٠	٦
١٠- ١٥	٨
١٥- ٢٠	١٦
٢٠- ٢٥	٦
المجموع	**٣٦**

إن رسم مضلع تكراري يعبر عن هذا التوزيع يتم من خلال رسم خط أفقي يمثل فئات المدة الزمنية وخط آخر رأسي يمثل عدد البرامج، مع مراعاة ترك مسافة (ولتكن نصف سنتيمتر) على جانبي الخط الأفقي؛ أي قبل بداية الفئة الأولى، وترك المساحة نفسها بعد نهاية الفئة الأخيرة. وفي منتصف كل فئة من الفئات، نتجه إلى أعلى ونضع نقطة تقابل تكرارات كل فئة على المحور الرأسي، ثم نقوم بتوصيل هذه النقاط فنحصل على المضلع التكراري، والذي يتخذ الشكل التالي:

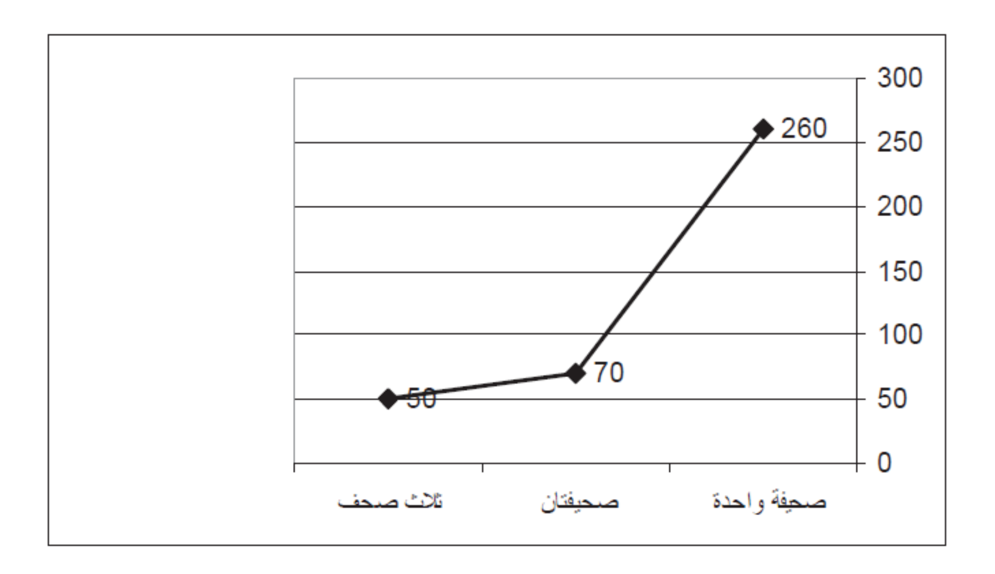


فكل نقطة (dot) من النقاط المبينة في هذا الرسم إنما توجد في منتصف الفئة، وكل نقطة أيضًا تقابل التكرارات الخاصة بالفئة، كما يلاحظ أنه تم إغلاق المضلع التكراري من الجانبين الأيسر والأيمن.

(د) المنحنيات:

المنحنيات (Curves): أداة مهمة في التعبير بوضوح عن البيانات، ويغطي استخدامها مدى واسعًا، ويمكن من خلالها التعبير بوضوح عن أكثر من متغير، ويجب وضع الإشارات المميزة لكل متغير، مع الالتزام بالدقة الشديدة عند رسمها بحيث تكشف عن مختلف أوجه الموضوع. لنفرض أن لدينا بيانات عن عينة قوامها ٤٠٠ شخص من قراء الصحف اليومية، موزعة حسب عدد الصحف اليومية التي يقرؤونها، وكان هذا التوزيع كالآتي:

صحيفة واحدة (٢٦٠)، صحيفتان (٧٠)، ثلاث صحف (٥٠)، في هذه الحالة يمكننا استخدام المنحنى لتمثيل هذا التوزيع، وذلك كالآتي:



فالمحور الأفقي يمثل عدد الصحف، أما المحور الرأسي فيمثل عدد القـ

توزيع العينة السابق ذكره، فهناك (٢٦٠) يقرؤون صحيفة يومية واحد

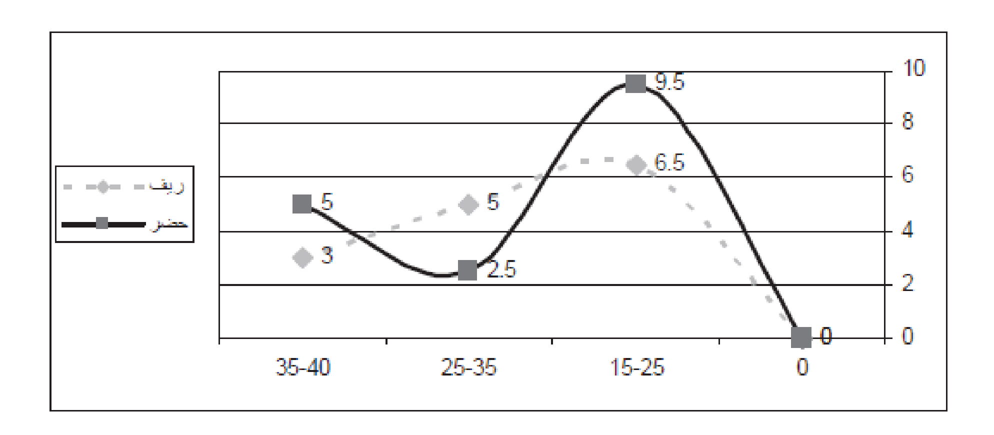
صحيفتين، أما الذين يقرؤون ثلاث صحف يومية فعددهم (٥٠).

كما يمكن استخدام المنحنيات في تمثيل أكثر من متغير تمثيلًا بيانيًا، لنفرض أن لدينا بيانات عن مشاهدة جمهور المجتمع المحلي للقنوات الفضائية التليفزيونية العامة، وتتوزع العينة حسب متغير فئات الأعمار ومتغير متوسط ساعات المشاهدة الأسبوعية على النحو الآتي:

فئات الأعمار	ريف	حضر
١٥-٢٥	٦.٥	٩.٥
٢٥-٣٥	٥	٢.٥
٣٥-٤٠	٣	٥

هنا يمكن التعبير عن هذه البيانات بالمنحنيات، وذلك برسم خط أفقي يمثل فئات الأعمار، وخط رأسي يمثل متوسط ساعات المشاهدة الأسبوعية، ثم نضع نقطة تمثل

ارتفاع كل فئة، فالفئة الأولى مثلًا (١٥-٢٥) نضع نقطة على ارتفاع ٦.٥ لتعبر عن متوسط وقت المشاهدة فيما يخص مجموعة الريف، كما نضع نقطة على ارتفاع ٩.٥ فيما يخص مجموعة الحضر، ونكرر هذه العملية في بقية الفئات، ثم نقوم بتوصيل النقاط التي تعبر عن متوسط وقت المشاهدة فيما يخص مجموعة الريف وذلك بخط معين، وكذلك لمجموعة الحضر ولكن بخط مختلف أو مميز عن خط مجموعة الريف، فيكون لدينا الرسم البياني التالي:

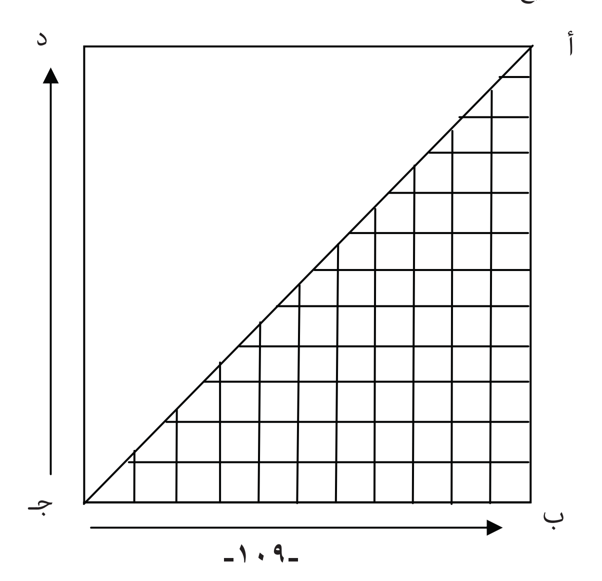


فالمحور الرأسي يوضح متوسط عدد ساعات المشاهدة الأسبوعية، والمحور الأفقي يوضح فئات الأعمار، ويكشف هنا المنحنى بوضوح عن أن:

- فيما يخص الفئة العمرية (١٥-٢٥)، يبلغ متوسط عدد ساعات المشاهدة الأسبوعية ٦.٥ ساعة في الريف مقابل ٩.٥ ساعة في الحضر.
- أما الفئة العمرية (٢٥ -٣٥)، فإن متوسط عدد ساعات المشاهدة الأسبوعية يبلغ ٥ ساعات للريفيين مقابل ٢.٥ ساعة للحضريين.

- وفيما يخص فئة الأعمار (٣٥- ٤٥)، يكشف المنحنى عن أن متوسط مشاهدة التليفزيون في مجموعة الريف يبلغ ثلاث ساعات في الأسبوع، بينما يصل إلى خمس ساعات فيما يخص مجموعة الحضر.

ويعتبر منحنى لورنز (Lorenz Curve) من الأشكال البيانية الهامة التي تجسد استفادة الدراسات العلمية من الرسوم البيانية لتوضيح الظواهر المعقدة، وقد توصل لورنز إلى هذا المنحنى في سياق توضيح مدى تركز الثروة (Wealth) لدى عدد قليل من الناس، غير أن تطبيقاته انتشرت فيما بعد في كافة تخصصات المعرفة بما فيها الاتصال، ويعتمد منحنى لورنز على المنحنى المتجمع الصاعد لنسب المتغيرين معًا، ولكل قيمة من قيم هذين المتغيرين نقاط التقاء على الوتر (أ د). ويتم تعيين نقاط إحداثيات المتغيرين على هذا الوتر، ويتحدد منحنى لورنز بمربع أضلاعه (أ، ب، ج، د) ، وهو مقسوم إلى مثلثين مشتركين في الوتر (أ د) هذا الوتر يمثل التوزيع الأمثل، وهو يعرف بخط التوزيع، فإذا كان توزيع الظاهرة يميل إلى يمين الوتر (أد)؛ أي باتجاه النقطة ب، فإن هذا الميل يعبر عن سوء التوزيع، وبقدر ما يكون توزيع الظاهرة في المثلث الأسفل أ، ب، ج يكون سوء توزيع هذه الظاهرة، فإذا زادت المساحة في هذا الجزء، دل ذلك على شدة سوء التوزيع، أما إذا تقلصت إلى أدنى حد أو اختفت، دل ذلك على جودة التوزيع.



ويستخدم منحنى لورنز في الكثير من التقارير والدراسات الإعلامية لوصف توزيع مرافق الاتصال بين المناطق المختلفة، وبين الفئات الاجتماعية، وكذلك توزيع محتوى الاتصال حسب معايير النوعية أو الجمهور المستهدف... إلخ.

(د) الدوائر:

كثيرًا ما تتفوق الدوائر على غيرها من وسائل العرض البياني؛ من حيث توضيح البيانات بكفاءة عالية، مثال ذلك توزيع أفراد المجتمع حسب امتلاك أجهزة الراديو والتليفزيون ، لنفرض أن مجمل سكان هذا المجتمع هو مليون شخص، وتم إحصاء السكان، وتبين أنهم يتوزعون حسب امتلاك أجهزة الراديو والتليفزيون كالآتي:

٧٠٠٠٠٠ يمتلكون جهاز راديو وجهاز تليفزيون

١٨٠٠٠٠ يمتلكون جهاز تليفزيون فقط

١٢٠٠٠٠ يمتلكون جهاز راديو فقط

في هذه الحالة يمكن تمثيل تلك المعطيات بالدائرة البيانية، وذلك بتقسيم الدائرة على أساس الزاوية المركزية ٣٦٠ درجة؛ بحيث يكون لدينا عدد من الزوايا مساويًا لعدد المجموعات المطلوب تمثيلها، وفي مثالنا المذكور فإن لدينا ثلاث مجموعات، والإجمالي هو مليون، فتكون:

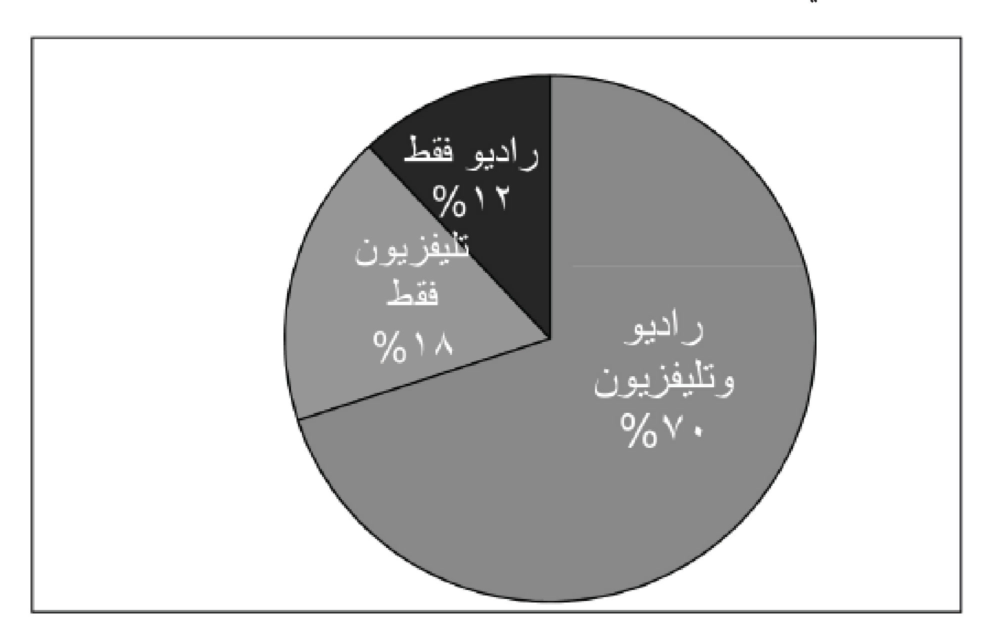
زاوية المجموعة الأولى هي: ٣٦٠ (٧٠٠٠٠٠ ÷ ١٠٠٠٠٠٠) = ٢٥٢

أما زاوية المجموعة الثانية فهي: ٣٦٠ (١٨٠٠٠٠ ÷ ١٠٠٠٠٠٠) = ٦٥

وتكون زاوية المجموعة الثالثة : ٣٦٠ (١٢٠٠٠٠ ÷ ١٠٠٠٠٠٠) = ٤٣

(يمكننا التأكد من صحة هذه العملية إذا كان إجمالي الزوايا ٣٦٠ درجة). هذا معناه أن المجموعة الأولى (التي تمتلك جهاز راديو وجهاز تليفزيون) تستغرق ٢٥٢ زاوية؛ أي ما

يعادل ٧٠% من زوايا الدائرة. أما المجموعة الثانية (التي تمتلك جهاز تليفزيون فقط) فتستغرق ٦٥ زاوية؛ أي ما يعادل ١٨% تقريبًا من زوايا الدائرة. وأخيرًا، فإن المجموعة الثالثة (التي تمتلك جهاز راديو فقط) تستغرق ٤٣ زاوية؛ أي ما يعادل ١٢% تقريبًا من زوايا الدائرة. هذه المعطيات تتضح في الدائرة البيانية كالآتي:



وبمجرد النظر إلى الدائرة، نتبين أن ٧٠% من أفراد المجتمع لديهم أجهزة راديو وتليفزيون، كما أن ١٨% لديهم أجهزة تليفزيون فقط، ولا يمتلكون أجهزة راديو، مقابل ١٢% لديهم أجهزة راديو فقط ولا يمتلكون أجهزة تليفزيون.

وتجدر الإشارة إلى أن برامج الحاسوب الإحصائية الجاهزة تمكن من تنفيذ جميع الأشكال البيانية البسيطة والمعقدة بسهولة ويسر،مع التأكيد على ضرورة أن يكون الرسم البياني سهلًا وبسيطًا؛ بحيث يدل مباشرة على الموضوع الذي يعبر عنه، وأن يستخدم الرموز الضرورية للتوضيح فقط، وأن يصحب الرسم تعليق سردي موجز يوضح الفكرة.

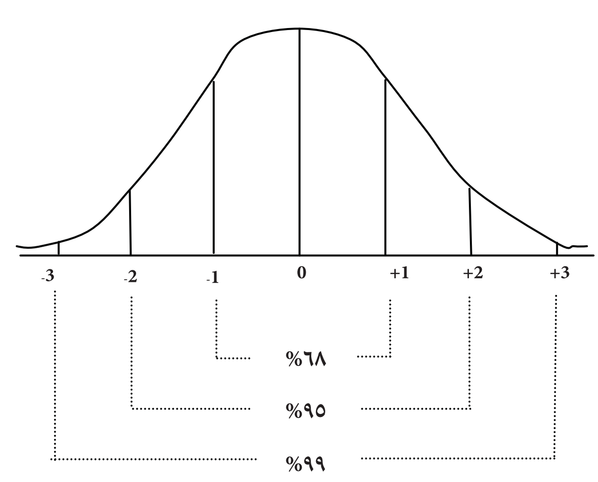
المبحث الثالث

وصف النزعة المركزية والتشتت

تتم المعالجة الإحصائية الوصفية باستخدام مقاييس النزعة المركزية ومقاييس التشتت، وعلى هذا الأساس، فإن المبحث الحالي يتضمن مطلبين، الأول يوضح مقاييس النزعة المركزية، أما الثاني فيوضح مقاييس التشتت.

المطلب الأول: مقاييس النزعة المركزية

النزعة المركزية (Central Tendency) هي نزوع التكرارات أو القيم الكميّة الخاصة بالموضوع إلى الاقتراب من أو الابتعاد عن القيمة الوسطى أو قيمة المركز، وهي القيمة التي تتجمع حولها أكثرية التكرارات، فإذا افترضنا أن هناك تكرارات توضح توزيع عينة عشوائية كبيرة من مشاهدي التليفزيون حسب عدد الساعات التي يقضونها يوميًا في المشاهدة، وكان متوسط ساعات المشاهدة ساعتين يوميًا، فإن غالبية العينة تكون قريبة من هذا (ساعتين) سواء بالزيادة أو بالنقصان، بينما يكون هناك عدد قليل يشاهد التليفزيون أقل من ساعتين، ويتناقص هذا العدد تدريجيًا كلما ابتعدنا عن الوسط الحسابي باتجاه ساعات المشاهدة الأقل، وهناك عدد قليل آخر يشاهد التليفزيون أكثر من ساعتين، وهذا العدد أيضًا يتناقص تدريجيًا كلما ابتعدنا عن الوسط الحسابي باتجاه ساعات المشاهدة الأكبر؛ أي أن العدد الأكبر من مفردات العينة يتجمع حول القيمة التي تعبر عن متوسط ساعات المشاهدة اليومية، ويقل هذا التجمع بالتدريج كلما ابتعدنا عن المتوسط يمينًا أو يسارًا، ويتم التعبير عن ذلك بما يعرف بالمنحنى الاعتدالي (Normal Curve) الذي يوضحه الرسم التالي:



فالعمود الرأسي النازل من منتصف قمة المنحنى إلى منتصف القاعدة يمثل القيمة الوسطى التي تتركز حولها معظم القيم، وهذا العمود يقسم قاعدة المنحنى إلى جزأين، كل جزء يتكون من ثلاث وحدات قيمة كل منها انحراف معياري واحد (الانحراف المعياري هو الانحراف عن الوسط الحسابي)، فالجزء الأيسر (-١، -٢، - ٣) انحراف معياري، أما الجزء الأيمن فهو (+١، +٢، + ٣) انحراف معياري، ومن خصائص المنحنى الاعتدالي أن:

٦٨% تقريبًا (وبالتحديد ٦٨.٢٦%) من القيم تقع بين (– ١) و (+ ١) انحراف معياري على يسار ويمين الوسط الحسابي

٩٥% تقريبًا (وبالتحديد ٩٥.٤٤%) من القيم تقع بين (– ٢) و (+ ٢) انحراف معياري على يسار ويمين الوسط الحسابي

٩٩% تقريبًا (وبالتحديد ٩٩.٧٤%) من القيم تقع بين (– ٣) و (+ ٣) انحراف معياري على يسار ويمين الوسط الحسابي

غير أن هذا التوزيع هو توزيع نظري مثالي، وفي الكثير من البحوث لا نحصل على توزيعات اعتدالية تمامًا، لكن معرفة شكل التوزيع يكون ضروريّا ليس فقط لوصف

البيانات بشكل مبسط وواضح، وإنما أيضا لتقرير استخدام أو عدم استخدام أساليب إحصائية معينة.في هذا الإطار تلعب مقاييس النزعة المركزية (Central Tendency) دورًا أساسيًا، ومن أبرز تلك المقاييس الوسط الحسابي (Arithmetic Mean)، الوسيط (Median)، المنوال (Mode)، الوسط الهندسي (Geometric Mean)، الوسط التوافقي (Harmonic Mean)؛ ونظرًا لأن المقاييس الثلاثة الأولى هي الأكثر شيوعًا وتناسب بحوث العلوم الاجتماعية، فسوف نشير إليها بإيجاز:

أولًا: الوسط الحسابي

وهو من أبسط مقاييس النزعة المركزية وأكثرها شيوعًا في الاستخدام، وتعتمد عليه مقاييس متعددة خاصة تلك التي ترصد وتحلل الفروق القائمة بين المجموعات في صفة معينة أو سلوك معين. وفي بحوث الاتصال يستخدم الوسط الحسابي على نطاق واسع لرصد وتحليل استجابات المفحوصين على المقاييس والاختبارات والاستبانات. فقد يكون هناك مقياس يقيس علاقة المشاهد بالتليفزيون أو اختبار يقيس المعرفة السياسية لدى مشاهدي التليفزيون أو مستمعي الراديو أو قراء الصحف.. ويتضمن المقياس أو الاختبار مجموعة من البنود التي لها استجابات تأخذ قيمًا كميّة معينة Quantitative values)، وباستخدام هذه القيم الكميّة يمكن الحصول على متوسط درجة العينة على المقياس أو الاختبار.

وهناك أيضًا متغيرات تجميعية، بمعنى تلك المتغيرات التي يتم إعدادها بحيث تشمل متغيرات فرعية كثيرة يتم تجميعها في متغير واحد يسمى (المتغير التجميعي)، مثال على ذلك متغير (المستوى الاقتصادي الاجتماعي)، وكذلك متغير (معدل التعرض لوسائل الإعلام)....، فمتغير المستوى الاقتصادي الاجتماعي مثلًا يتضمن مؤشرات عديدة مثل: الدخل، التعليم ، منطقة السكن ، الاشتراك في نادي ثقافي، امتلاك الأجهزة الحديثة ... إلخ، كما أن متغير معدل التعرض لوسائل الإعلام يشمل الوقت الذي يقضيه الفرد في سماع الإذاعة، مشاهدة التليفزيون، قراءة الصحف، استخدام الإنترنت... إلخ، ويتم

(تجميع) القيم الكميّة لاستجابات المبحوثين على بنود أو جوانب مقياس المستوى الاقتصادي الاجتماعي؛ ومن ثمّ نحصل على متوسط درجة هؤلاء المبحوثين على هذا المقياس، ولذلك يتم تصنيفهم حسب المستوى الاقتصادي الاجتماعي.. المنطق نفسه فيما يخص معدل التعرض لوسائل الإعلام.

وكثيرًا ما يتطلب البحث الإعلامي معرفة معنوية الفروق بين المفحوصين؛ من حيث استخدام وسائل الإعلام حسب المستوى الاقتصادي الاجتماعي (ككل وليس حسب كل مؤشر على حدة)، وفي هذا الحالة يتم تصنيف العينة حسب المستوى الاقتصادي الاجتماعي (منخفض، متوسط، مرتفع)، ويتم استخراج الوسط الحسابي لاستجابات العينة على مقياس استخدام وسائل الإعلام؛ ومن ثمّ يمكن المقارنة بين مجموعات العينة ومعرفة ما إذا كانت بينها فروق جوهرية في استخدام وسائل الإعلام حسب المستوى الاقتصادي الاجتماعي، كما أن بحوث الإعلام كثيرًا ما تستخدم مقاييس تتضمن بنودًا متعددة لقياس موضوع واحد لمعرفة متوسط درجة المفحوصين على المقياس ككل، وبالتالي تبدو ضرورة رصد وتحليل متوسط مجمل استجابات العينة على تلك البنود مجتمعة وليس على كل بند بمفرده. كمثال على ذلك نفرض أن لدينا عينة قوامها ٤٠٠ مفردة وأجابوا على مقياس يتضمن عشرين برنامجًا تليفزيونيّا، كان المطلوب من كل مفحوص أن يختار إجابة واحدة من الإجابات الثلاث الآتية بشأن كل برنامج:

- يشاهد بصفة منتظمة.
- يشاهد بصفة غير منتظمة.
- لا يشاهد.

وعند تكويد أو ترميز تلك الاستجابات، يراعى أن تتخذ هذه الاستجابات الثلاث أرقامًا متدرجة تنازلية (٣، ٢، ١) على التوالي، وليس (١، ٢، ٣).. المنطق في ذلك أن المشاهدة بصفة منتظمة تكون أكبر من المشاهدة بصفة غير منتظمة، كما أن المشاهدة غير

المنتظمة تفوق عدم المشاهدة، وبالتالي لا بد أن تعكس الأرقام هذه الحقيقة، بمعنى أن الرقم الدال على المشاهدة المنتظمة يكون أكبر من الرقم الدال على المشاهدة غير المنتظمة، كما أن الرقم الدال على المشاهدة غير المنتظمة لا بد أن يكون أكبر من الرقم الدال على عدم المشاهدة، وبدون مراعاة ذلك، فإن النتائج تأتي مضللة تمامًا (وهذا الخطأ يقع فيه الكثير من الباحثين).

إن الأرقام (3، 2، 1) هي التي يتعامل معها الإحصاء، وبناء على تلك الأرقام، فإن الدرجة التي تعكس استجابات المفحوصين على المقياس المذكور تتراوح ما بين (20) إلى (60)؛ وذلك لأن المقياس يتضمن عشرين بندًا، كل بند يعبر عن برنامج تليفزيوني معين؛ أي أن كل مفحوص سيحصل على درجة لا تقل عن 20 ولا تزيد على 60، ولكي نفهم الوسط الحسابي على نحو أفضل، نوضح فيما يلي أهم طرق حسابه:

(أ) حساب الوسط الحسابي من البيانات غير المبوبة:

إذا كان لدى الباحث قيم كميّة عبارة عن بيانات غير مبوبة في جدول، فإن الوسط الحسابي هو حاصل مجموع هذه القيم مقسومًا على عددها.إنه ناتج قسمة المجموع الجبري لهذه القيم، ومعادلته هي:

$$م = \frac{١}{ن} (س1 + س2 + س3 + س 4+....س ن)$$

$$= \frac{١}{ن} (مجـ س)$$

$$= \frac{مجـ س}{ن}$$

حيث إن:

م : هو الوسط الحسابي

ن: عدد الأفراد (حجم العينة)

س: قيم المفردات

مجـ س: إجمالي قيم المفردات

فـــإذا كـــان لـــدينا ثلاثـــة أعـــداد هـــي : 38،54،26 فـــإن الوســـط الحســـابي يســـاوي $38 + 54 + 26 \div 3 = 39.3$

وإذا كان لدينا عينة من 500 شخص وطبق عليهم مقياس معين، فإن كل شخص سيحصل على درجة معينة حسب استجاباته على بنود المقياس، فإذا جمعنا درجات جميع الأشخاص وقسمنا المجموع على 500، يكون الناتج هو الوسط الحسابي؛ أي متوسط درجة استجابة العينة على المقياس.

(ب) حساب الوسط الحسابي من تكرارات الأعداد المنفصلة:

كثيرًا ما يكون لدى الباحث بيانات عبارة عن مجموعة من الأعداد المنفصلة (تمثل حالات معينة)، ولكل عدد أو حالة تكرار معين، وهنا يمكن الحصول على الوسط الحسابي من خلال ضرب كل عدد في التكرارات الخاصة به، ثم جمع الناتج وقسمته على عدد الحالات، مثال ذلك جدول تكراري يوضح عدد الإعلانات (ن) وأمام كل منها عدد القراء الذين أطلعوا عليه (ك). إن الوسط الحسابي لهذا الجدول يتم من خلال ضرب عدد الإعلانات في عدد القراء المقابل له، ثم نجمع الناتج ونقسمه على عدد الإعلانات؛أي أن المعادلة المستخدمة في هذه الحالة هي:

$$\text{م} = \frac{\text{مجـ (س ك)}}{\text{ن}}$$

حيث: م: الوسط الحسابي

س: قيمة الفئة

ك: تكرار الفئة ، ن: عدد الحالات

مثال: في دراسة استخدمت مجموعة النقاش المركزة حول قراءة الإعلانات المنشورة بالصحف اليومية، حصلت الباحثة على النتيجة المبينة بالجدول الآتي:

عدد الإعلانات	عدد القراء
٥	١٣
٧	١٤
٨	١٢
٩	١٥
١٠	١١

إن إجمالي عدد الإعلانات هو (٣٩) إعلانًا؛ أي أن قيمة (ن) = ٣٩ وتتوزع هذه الإعلانات حسب عدد القراء الذين اطلعوا عليها، فهناك خمسة إعلانات بلغ عدد الذين قرؤوها (١٣) شخصًا، وهناك سبعة إعلانات بلغ عدد الذين قرؤوها (١٤) شخصًا.. وهكذا، فإذا رمزنا إلى عدد الإعلانات بالرمز (س)، وإلى التكرارات (عدد القراء) بالرمز (ك)، فإنه عند ضرب كل عدد من الإعلانات في التكرارات المناظرة له نحصل على الجدول الآتي:

عدد الإعلانات (س)	عدد القراء (ك)	س × ك
٥	١٣	٦٥
٧	١٤	٩٨
٨	١٢	٩٦
٩	١٥	١٣٥
١٠	١١	١١٠
المجمـــــوع		٥٠٤

إن مجموع (س × ك) هو (٥٠٤)، وبما أن عدد الإعلانات (ن) = ٣٩، فيكون الوسط الحسابي: ٥٠٤ ÷ ٣٩ = ١٢.٩٢؛ أي ١٣ تقريبًا

(ب) حساب الوسط الحسابي من جدول الفئات التكرارية:

فالجدول التكراري ذو الفئات يتضمن تكرارات موزعة على فئات معينة، لكل فئة حد أدنى وحد أعلى؛ أي أن لكل فئة طول أو مدى معين، ولحساب الوسط الحسابي في هذه الحالة يتم تحديد منتصف كل فئة (ص) وضربه في تكرارات الفئة نفسها،ثم نجمع الناتج ونقسمه على إجمالي عدد التكرارات (ن). فإذا افترضنا أن الفئة الأولى (٣٠-٣٩)، فإن منتصفها يكون: ٣٠+٣٩÷٢ = ٣٤.٥ فإذا كان عدد تكرارات هذه الفئة هو (٨) مثلًا، فإن قيمة (ص ك) تساوي ٣٤.٥× ٨ = ٢٧٦ وبهذه الطريقة نحصل على قيمة (ص ك) لكل فئة، وبجمع (ص ك) لكل الفئات نحصل على (مج ص ك)، وباستخدام هذه المعطيات نحصل على الوسط الحسابي بموجب التعويض في المعادلة:

$$م = \frac{مجـ \ (ص\ ك)}{ن}$$

حيث: ص: منتصفات الفئات

ك: تكرارات الفئات ، ن: مجمل التكرارات

مثال: الجدول الآتي يوضح فئات المساحة بالسم[٢] لعدد من الأخبار المنشورة بجريدة يومية:

فئات المساحة	ك
٣٠-٣٩	٨
٤٠-٤٩	٤
٥٠-٥٩	٦
٦٠-٦٩	١٠
٧٠-٨٠	٤

بتطبيق الخطوات السابقة يتم حساب منتصف كل فئة (ص)، وذلك بجمع الحد الأدنى والحد الأعلى لكل فئة، ثم قسمة الناتج على ٢، فالفئة الأولى حدها الأدنى (٣٠) وحدها الأعلى (٣٩) وبالتالي يكون منتصفها ٣٠+٣٩÷٢ = ٣٤.٥، وهكذا فيما يخص

جميع الفئات ثم نضرب منتصف الفئة في عدد تكراراتها، فنحصل على (ص ك) وفق ما يوضحه الجدول الآتي:

الفئات	ك	ص	ص × ك
٣٠- ٣٩	٨	٣٤.٥	٢٧٦
٤٠- ٤٩	٤	٤٤.٥	١٧٨
٥٠- ٥٩	٦	٥٤.٥	٣٢٧
٦٠- ٦٩	١٠	٦٤.٥	٦٤٥
٧٠- ٧٩	٤	٧٤.٥	٢٩٨
المجموع	**٣٢**		**١٧٢٤**

أي أن مج ك هو (٣٢)، كما أن مج ص ك هو (١٧٢٤)، وبقسمة هذا الناتج على عدد التكرارات نحصل على الوسط الحسابي؛ أي أن:

$$م = \frac{1724}{32} = ٥٤$$

(جـ) حساب الوسط الحسابي باستخدام وسط فرضي:

تستخدم هذه الطريقة عندما تكون الفئات متساوية المدى أو الطول. والوسط الفرضي هو قيمة ضمن فئة من فئات الجدول، يختارها الباحث كنقطة توسط أو نزعة مركزية وسطى لجميع الفئات، ومن هذه الفئة يبدأ تدريج الفئات بالزيادة في اتجاه معين والتناقص في الاتجاه الآخر. ويتم تحديد الوسط الفرضي بطريقة بسيطة جدّا. فإذا افترضنا أن لدينا جدولًا يتضمن توزيعًا تكراريًا على فئات معينة (ولتكن خمس فئات مثلًا)، فإن الباحث يختار فئة المنتصف أو أقرب فئة إلى المنتصف، كأن تكون الفئة الثالثة أو الفئة الرابعة، لنفرض أننا اخترنا الفئة الثالثة وكان مداها (٣٠ إلى ٤٠)، إن الوسط الفرضي عبارة عن ٣٠+٤٠÷٢ = ٣٥؛ أي أن قيمة الوسط الفرضي هي (٣٥) هذا الوسط الفرضي نطرحه (ننقصه) من منتصف كل فئة (ص)، ويتم الحصول على منتصف كل فئة بأن نجمع قيمة بداية الفئة + قيمة نهايتها ثم القسمة على (٢)، وبالحصول على الوسط الفرضي للفئة

الوسطى، يتم إنقاصه من منتصف كل فئة من فئات الجدول، فالفئة الأولى بالجدول مثلًا قد تكون (10 إلى 20)، وبالتالي منتصف هذه الفئة يكون (10+ 20÷ 2 = 15)، فإذا كانت قيمة الوسط الفرضي (35)، فإن قيمة (ص) للفئة الأولى تكون 15- 35 = -20 وهكذا في كل فئات الجدول، وبهذا يكون لدى الباحث المعطيات الأساسية لحساب الوسط الحسابي باستخدام المعادلة:

$$م = ض + \left(\frac{(مجـ\ ص\ ك)}{ن} \right)$$

حيث: م: الوسط الحسابي

ض : الوسط الفرضي

ص ك : منتصف كل فئة مضروبًا في عدد تكراراتها

مجـ ص ك: مجموع ص ك لكل الفئات

ن: عدد الحالات

مثال توضيحي: نفرض أننا نريد الحصول على الوسط الحسابي باستخدام طريقة الوسط الفرضي من التوزيع التكراري السابق. نلاحظ أن الجدول يتضمن خمس فئات، وأن الفئة الوسطى هي الفئة الثالثة (50 إلى 59). إن الوسط الفرضي = 50 + 59 ÷ 2 = 54,5، بإنقاص هذه القيمة من منتصفات الفئات نحصل على (الفرق) بين الوسط الفرضي ومنتصفات جميع الفئات، ففي الجدول الآتي نجد أن الفئة الأولى مثلًا هي (30 إلى 39)، وبالتالي، فإن منتصفها يكون (30+ 39÷2= 34.5)؛ ونظرًا لأن الوسط الفرضي يساوي 54,5، فإن (ص) لهذه الفئة تساوي 34.54 – 54.5 = -20 وهكذا فيما يخص جميع الفئات على النحو الذي يوضحه الجدول الآتي:

الفئات	ك	منتصف الفئات	ص	ص × ك
39-30	8	34.5	20- = 54.5 – 34.5	160-
49-40	4	44.5	10- = 54.5 – 44.5	40-
59-50	6	54.5	صفر-54.5 = صفر	صفر
69-60	10	64.5	10+ = 54.5 – 64.5	100
79-70	4	74.5	20+ = 54.5 – 74.5	80
المجموع	32			= -20

أي أن مج ص ك = - 20 كما أن الوسط الفرضي = 54.5 أما عدد التكرارات فهو (32)؛ أي أن قيمة (ن) تساوي (32)، ويتم الحصول على الوسط الحسابي باستخدام المعادلة:

$$م = ض + \left(\frac{مج\ ص\ ك}{ن}\right)$$

$$\therefore\ م = 54.5 + \left(\frac{-٢٠}{٣٢}\right) = 53.875$$

(د) حساب الوسط الحسابي باستخدام الانحرافات البسيطة:

وهي نفس الطريقة السابقة ولكن مع قسمة قيمة (ص) على طول الفئة، ويتم حساب الوسط الحسابي بموجب المعادلة:

$$م = ض + \left(\frac{مج\ ص\ ك}{ن} \times ط\right)$$

حيث: م: الوسط الحسابي

ض: الوسط الفرضي

ص ك: منتصف كل فئة مضروبًا في عدد تكراراتها

مجـ ص ك: مجموع ص ك

ن: عدد الحالات

ط : طول الفئة

وبالعودة إلى الجدول السابق، نجد أن الفئة الأولى مثلًا هي (30 - 39)؛ أي أنها تمتد من (30) إلى (39)، وبالتالي فإن طولها يساوي (10)، وبالنظر إلى قيمة (ص) لهذه الفئة في الجدول السابق نجدها (-20)، فإذا قسمنا هذه القيمة على طول الفئة؛ أي على (10) ، فإن الناتج = -20 ÷ 10 = -2 وهكذا في جميع الفئات، فنحصل على قيم (ص) بطريقة الانحرافات البسيطة، على النحو المبين بالجدول التالي:

الفئات	ك	ص	ص × ك
39-30	8	-20 ÷ 10 = -2	-16
49-40	4	-10 ÷ 10 = -1	-4
59-50	6	صفر = صفر	صفر
69-60	10	+10 ÷ 10 = 1	10
79-70	4	+20 ÷ 10 = 2	8
المجموع	**32**		**-2**

وكما هو واضح، فإن مج ص ك هو (-2)، كما أن الوسط الفرضي = 54.5 أما عدد التكرارات فهو (32)، وبالعودة إلى معادلة حساب الوسط الحسابي بطريقة الانحرافات البسيطة، فإنها كما سبقت الإشارة هي:

$$\text{م} = \text{ض} + \left(\frac{\text{مجـ ص ك}}{\text{ن}} \times \text{ط} \right)$$

وبالتعويض في هذه المعادلة يكون:

$$\text{م} = 54.5 + \left(\frac{-٢}{٣٢} \times ١٠ \right)$$

= ٥٤.٥ – ٠.٦٢٥ = ٥٣.٨٧٥

ويلاحظ أننا قسمنا (مج ك ص) على (ن) مع الضرب في الرقم (١٠) الذي هو طول الفئة.

وتتيح الحزمة الإحصائية للعلوم الاجتماعية (SPSS) هذه العملية بسهولة، وذلك من خلال إجراءات معينة يتم تنفيذها بواسطة الحاسوب بموجب أوامر محددة، فنحصل على الوسط الحسابي. افترض أن هناك دراسة أجريت على عينة قوامها ٤٠٠ مفردة، لمعرفة استخدامات التليفزيون (TV Using)، وقد تم تطبيق مقياس مقنن على هذه العينة، وكانت الدرجة على هذا المقياس تتراوح ما بين (٢٠) إلى (٤٠) درجة، وتم إدخال بيانات تلك الدراسة في الحاسوب وعولجت إحصائيًا باستخدام برنامج (SPSS) بما في ذلك حساب مجمل استجابات كل مفحوص على مجمل بنود المقياس ، فإننا نحصل على المعطيات الآتية:

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Tvusing	400	2.00	36.00	23.8175	4.78607
Valid N (listwise)	400				

يتضمن هذا الجدول الوسط الحسابي والانحراف المعياري، والمدى (أقل درجة وأعلى درجة)، ومنه نتبين أن :

م (الوسط الحسابي) = ٢٣.٨٢

ع (الانحراف المعياري) = ٤.٨

أقل درجة هي ٢ وأعلى درجة هي ٣٦

هكذا يمكن القول إن متوسط درجة العينة (ن=٤٠٠) على مقياس استخدام التليفزيون هو ٢٣.٨٢ درجة بانحراف معياري ٤.٨، وقد تراوحت درجات المفحوصين ما بين ٢ إلى ٣٦، ويتضح من ذلك أن متوسط درجة المفحوصين يعادل حوالي ٦٠% من الدرجة العظمى للمقياس والتي تبلغ ٤٠ درجة كما سبقت الإشارة. وبالمثل ومن خلال الحزمة الإحصائية للعلوم الاجتماعية (SPSS) يمكن الحصول على متوسطات درجات جميع المجموعات الفرعية التي تضمها العينة أيّا كان المتغير أساس التصنيف، كمتغير الجنس (ذكور & إناث)، وكذلك متغير محل الإقامة (ريف & حضر)، ومتغير السن (أقل من ٣٠ سنة & ٣٠ سنة فأكثر) ... إلخ، ويمكن المقارنة بين متوسطات درجات تلك المجموعات لمعرفة معنوية الفروق بينها من حيث الدرجة على مقياس استخدام التليفزيون.

خصائص الوسط الحسابي:

١) إن مجموع انحرافات القيم عن وسطها الحسابي يساوي الصفر . فإذا كان لدينا ثلاث قيم هي: ١٢ ، ١٣ ، ١٤، فإن متوسط هذه القيم هو ١٢+١٣ + ١٤= ٣٩ ÷ ٣= ١٣

أي أن متوسط القيم المذكورة هو ١٣ وعند حساب انحراف كل قيمة عن الوسط نجد أن:

- انحراف القيمة الأولى= ١٢-١٣= -١
- انحراف لقيمة الثانية = ١٣-١٣= صفر
- انحراف القيمة الثالثة = ١٤-١٣= ١

فإذا حسبنا مجمل تلك الانحرافات، نجد أنه يساوي الصفر:

(-١+ صفر +١= صفر)

٢) إذا كان هناك وسطان حسابيان (سَ ، صَ) لسلسلتين من البيانات، فإن الوسط الحسابي لمجموع قيمهما يساوي مجموع هذين الوسطين، بمعنى أن:

جَ = س- + صَ

حيث إن جَ هي الوسط الحسابي لمجموع قيم الوسطين، وعليه فإن:

ج١= س١ + ص١

ج٢= س٢ + ص٢

ج٣= س٣ + ص٣

ج ن= س ن + ص ن

جَ = سَ + صَ

٣) إذا كان لدينا سلسلة (ن) من المفردات للمتغير (س)، ولتكن مشاهدة التليفزيون في اليوم (نصف ساعة، ساعة) وهذه السلسلة تتكون من بعدين فرعيين (ذكور وإناث)، فإن الوسط الحسابي لقيم هذين البعدين الفرعيين يمكن أن يكون هو متوسط قيم السلسلة (ن)، فإذا كانت قيمة البعد الأول (الذكور) هي (س١) وكانت قيمة البعد الثاني (الإناث) هي (س٢)، فإن النتيجة هي:

س- = س-١ + س-٢

٤) إذا أعطينا قيمة الوسط الحسابي لكل مفردة من المفردات، فإن المجموع الناتج يساوي مجموع عدد المفردات. مثال: لدينا خمسة أرقام هي: ٢، ٤، ٦، ٨، ١٠

الوسط هو ٢ + ٤ + ٦ + ٨ + ١٠ = ٣٠ ÷ ٥ = ٦

أي أن متوسط هذه الأرقام يساوي (٦) كما أن مجموعها يساوي (٣٠)، فإذا وضعنا الرقم ٦ بدل كل رقم من الأرقام الخمسة المذكورة، فإن الناتج يكون ٦ + ٦ + ٦ + ٦ + ٦ = ٣٠ وهو نفس مجموع هذه الأرقام

٥) يتأثر الوسط الحسابي بالقيم المتطرفة (القيم الكبيرة جدّا والقيم الصغيرة جدًا). فإذا كان هناك شخصان، الأول يشاهد عشرين فيلمًا سينمائيًا في العام، أما الثاني فيشاهد أربعة أفلام ، فإن متوسط عدد الأفلام هو ١٢ فيلمًا (٢٠+ ٤÷٢= ١٢)، وهذه

القيمة للمتوسـط تبتعـد كثيرًا عن عدد الأفلام التي يشاهدها كل شخص، هذا الابتعاد عن القيم الحقيقيـة ناتج عن تأثر الوسط الحسابي بالقيم المتطرفة فهناك القيمـة ٢٠ وهـي قيمـة متطرفـة (كبيـرة جدّا) مقارنة بالقيمة ٤ وهي قيمة متطرفة (صغيرة جدّا). الوضع يختلف عندما لا تكـون قيم متطرفة، فإذا كان هناك شخصان، الأول يشاهد (١٤) عشر فيلمًا سينمائيّا في العام، أمـا الثـاني فيشاهد (١٢) فيلمًا، فإن الوسط يكون (١٣) فيلمًا، وهذه القيمة للمتوسط تقترب كثيرًا من عدد الأفلام التي يشاهدها كل شخص، هذا الاقتراب من القيم الحقيقيـة نـاتج عـن عـدم وجـود قـيم متطرفة؛ إذ يتقارب الشخصان من حيث عدد الأفلام السينمائية التي يشاهدانها خلال العام.

٦) يتعذر إيجاد الوسط الحسابي من الفئـات التكراريـة المفتوحـة، فـإذا كـان لـدينا توزيـع تكـراري يتضمن مثل هذه الفئات، يكون من الضروري إما إغلاقها،أو استبعادها، وحساب الوسط الحسابي لبقية الفئات، أو أن نطبق قوانين إحصائية معينة للحصول على قيمة تقريبية للوسط الحـسابي، أو استخراج الوسط الحسابي بالرسم الهندسي.

٧) عنـدما نحـصل عـلى الوسـط الحـسابي لعـدد مـن القـيم، فـإن قيمـة الوسـط الحـسابي قـد لا تكون متضمنة في تلك القيم. على سبيل المثـال، فإن الأعداد ٤، ٥، ٩ متوسطها الحسابي هـو ٦؛ إذ إن (٤ + ٥ + ٩ ÷ ٣ = ٦) أي أن قيمة الوسط الحسابي <u>غـير موجـودة</u> في القـيم الفعليـة، لكـن هذا ليس دائمًا، فالأعداد ١٢، ١٣، ١٤ متوسطها الحسابي هـو ١٣؛ إذ إن (١٢+ ١٣+ ١٤ ÷ ٣= ١٣) أي أن قيمة الوسط موجودة بين القيم الفعلية.

٨) ثبت أن الوسط الحسابي لا يصلح كمقياس للنزعة المركزية في بعض الحالات أهمها:

- عندما يكون شكل التوزيع على شكل الحرف U أو ما يعرف بالإنجليزية(U- Shaped Distribution) (يمكن التحقق من ذلك).

- عندما تكون البيانات في صورة متوالية هندسية مثال: (2، 4، 8، 16، 32، 64 ... إلخ)، ففي هذه الحالة يفضل استخدام الوسط الهندسي (Geometric Mean).

- عند حساب معدلات تغير الظاهرة عبر فترة زمنية معينة، ومعدلات تسارع وتيرة الظاهرة.

وعلى الرغم من ذلك فإن الوسط الحسابي يعتبر ضمن المعاملات الإحصائية شائعة الاستخدام لما فيه من جوانب قوة، فهو مثلًا يأخذ في الاعتبار جميع مفردات العينة، كما أنه أقل تأثرًا بالتغيرات التي تطرأ على حجم العينة، ويمكن معالجته جبريًا، ولا تتأثر قيمة الوسط الحسابي كثيرًا عند إعادة تنظيم الفئات، كدمج فئتين مع بعضهما، أو إعادة توزيع المفردات على فئات أكثر، كما أن الوسط الحسابي - بالإضافة إلى سهولته- يمكن حسابه من فئات غير متساوية، وذلك بإيجاد مركز كل فئة، وضرب مركز كل فئة في تكراراتها.

ثانيًا: الوسيط

الوسيط (Median) هو القيمة التي تتوسط القيم؛ بحيث يكون عدد القيم الموجودة على يسار تلك القيمة مساويًا لعدد القيم الموجودة على يمينها، ويمكن حساب الوسيط من البيانات الخام مباشرة، كما يمكن حسابه من البيانات المبوبة، وكذلك يمكن حسابه باستخدام المنحنى التكراري لأي توزيع على النحو التالي:

(أ) حساب الوسيط من البيانات غير المبوبة:

افترض أن لدينا هذه القيم الخمس: 6، 5، 4، 8، 7 فإذا قمنا بترتيب هذه القيم تصاعديًا ، فإنها تكون: 4، 5، 6، 7، 8 وهنا نلاحظ أن القيمة 6 هي التي تتوسط هذه القيم، فعلى يمين القيمة 6 توجد قيمتان (هما القيمة 4 والقيمة 5) وعلى يسارها توجد أيضًا قيمتان (هما القيمة 7 والقيمة 8)؛ أي أن عدد القيم الموجودة على (يمين) الرقم 6 يساوي عدد القيم الموجودة على(يساره) وبالتالي، فإن القيمة 6 هي الوسيط. هذا يعني أنه

عندما نقوم بترتيب مجموعة من الأعداد ترتيبًا تصاعديًا أو تنازليًا، فإن العدد الذي يتوسط تلك الأعداد يكون هو الوسيط، وهذا ينطبق فقط على الأعداد غير المبوبة التي إذا أحصيناها، تعطي عددًا فرديًا (كأن تكون ثلاثة أعداد، أو خمسة أعداد، أو سبعة أعداد ... إلى آخر الأعداد الفردية)، لكن الأعداد، أو المفردات قد يكون عددها زوجيًا، فمثلًا إذا كان لدينا الأعداد:

9، 6، 5، 4، 8، 7

أي أنها تعطي عددًا زوجيًا (ستة أعداد)، فلا يوجد عدد يتوسط هذه القيم تمامًا، في هذه الحالة، يتم الحصول على الوسيط من خلال ترتيب هذه القيم تنازليًا أو تصاعديًا (4، 5، 6، 7، 8، 9)، ثم جمع القيمتين اللتين تتوسطان بقية القيم، ثم قسمة هاتين القيمتين على 2 فيكون الناتج هو الوسيط، ويلاحظ أن القيمة 6 والقيمة 7 هما القيمتان اللتان تتوسطان بقية القيم، فعلى اليمين منهما توجد قيمتان وعلى اليسار منهما توجد قيمتان، وعلى هذا الأساس فإن الوسيط يكون:

$$6 + 7 \div 2 = 6.5$$

وهناك قانون للحصول على الوسيط من بيانات غير المبوبة بوجه عام، سواء كانت بيانات تعطي عددًا زوجيًا أو عددًا فرديًا، ويرتكز هذا القانون على تحديد ترتيب الوسيط؛ ومن ثمّ حسابه بموجب المعادلة:

$$\text{ت . و} = \frac{\text{ن} + ١}{٢}$$

حيث إن :

ت. و : ترتيب الوسيط

ن : عدد القيم

مثال: فيما يلي عدد المسلسلات التليفزيونية التي شاهدها سبعة أشخاص، وكذلك الكتب التي قرؤوها خلال عام، والمطلوب إيجاد الوسيط لكل من عدد المسلسلات وعدد الكتب:

الشخص	عدد المسلسلات	عدد الكتب
الأول	4	1
الثاني	2	3
الثالث	5	2
الرابع	3	4
الخامس	7	6
السادس	9	5
السابع	8	--

ففيما يخص المسلسلات يلاحظ أن لدينا 7 قيم (أي عدد فردي)، ونبدأ بترتيب عدد المسلسلات ترتيبًا تصاعديًا (من الأصغر إلى الأكبر)، فيكون: 2، 3، 4، 5، 7، 8، 9 وبتطبيق القانون المشار إليه يكون:

$$\text{الوسيط} = \frac{\text{ن} + ١}{٢}$$

بما أن لدينا سبعة أعداد (أي أن قيمة ن = 7)، فيكون ترتيب الوسيط يساوي:

$$\frac{٧ + ١}{٢} = 4$$

أي أن ترتيب الوسيط هو الترتيب الرابع، وتكون القيمة المقابلة لهذا الترتيب هي الوسيط، فإذا نظرنا إلى الترتيب التصاعدي للمسلسلات نجد أن القيمة التي تقع في الترتيب الرابع هي (5)؛ أي أن الوسيط للأعداد 2، 3، 4، 5، 7، 8، 9 يساوي (5)، هذا فيما يخص المسلسلات.

أما فيما يخص الكتب، يلاحظ أن لدينا 6 ستة أعداد؛ أي أن قيمة ن = 6 ، وتتم الخطوات السابقة نفسها، بمعنى ترتيب أعداد الكتب ترتيبًا تصاعديًا (من الأصغر إلى الأكبر)، فيكون: 1، 2، 3، 4، 5، 6

بتطبيق القانون المشار إليه يكون:

$$\text{ترتيب الوسيط} = \frac{\text{ن}+١}{٢}$$

بما أن قيمة (ن) هي 6 إذن ترتيب الوسيط يساوي

$$\frac{٦+١}{٢} = 3.5$$

أي أن ترتيب الوسيط هو (3.5) وبالنظر إلى الأعداد 1، 2، 3، 4، 5، 6 فإن هذا الترتيب يقع بين القيمة 3 والقيمة 4، ومن ثمّ يكون الوسيط:

3 + 4 ÷ 2 = 3.5 وعلى هذا الأساس فإنه في حالة الأعداد الفردية يكون ترتيب الوسيط هو قيمته.

وفي حالة وجود عدد صغير من القيم، يمكن حساب الوسيط مباشرة بعد ترتيبها، ففي الأعداد 1، 2، 3، 4 ، 5 ،6 نجد أن القيمتين (3) و(4) تتوسطان هذه الأعداد، وبجمع هاتين القيمتين ثم قسمة الناتج على 2 نحصل على الوسيط، غير أن الأمر يختلف في الواقع العملي؛ حيث يكون عدد القيم بالمئات أو الألوف مثلًا، ففي هذه الحالة يستغرق ترتيب القيم وقتًا وجهدًا، وتزداد احتمالات الخطأ، ومن هنا يكون استخدام الحاسب الآلي أمرًا ضروريًا.

(ب) حساب الوسيط من تكرارات الأعداد المنفصلة:

إذا كان لدينا مجموعة من القيم الكميّة (الأعداد)، ولكل قيمة تكرار معين، فإن أسهل طريقة لحساب الوسيط هي استخدام المنحنى التكراري المتجمع الصاعد. ولتوضيح هذه الطريقة نقدم هذا الجدول الذي يوضح المدة الزمنية بـ(الدقيقة) لعدد حلقات برامج إذاعية:

المدة بالدقيقة	التكرارات
١٠	٤
١٢	٥
١٤	٨
١٦	٦
١٨	٣
المجموع	**٢٦**

فالتكرارات في هذا الجدول، توضح أن هناك (٤) حلقات مدة كل منها عشر دقائق، وهناك (٥) حلقات مدة كلٍّ منها اثنتا عشرة دقيقة.. وهكذا.

وبتكوين المنحنى التكراري المتجمع الصاعد يكون لدينا الجدول التكراري الآتي:

المدة بالدقيقة	التكرارات	التكرار المتجمع الصاعد
١٠	٤	٤
١٢	٥	٩
١٤	٨	١٧
١٦	٦	٢٣
١٨	٣	٢٦
المجموع	**٢٦**	

وبما أن عدد التكرارات هو (٢٦)، فإن رتبة الوسيط = ٢٦÷٢= ١٣

وإذا كانت رتبة الوسيط تساوي (١٣)، فإن هذه القيمة تقع بين الرقم (٩) والرقم (١٧) في التوزيع التكراري المتجمع الصاعد، كما أن القيمة (١٣) تعني أن الوسيط يقع بين

الدرجة (12) والدرجة (14) في فئات الوقت بالدقيقة، وهذا يعني أن الزيادة بين الوسيط والدرجة 12 تناظر الزيادة بين رتبة الوسيط (13) والقيمة (9)، ويتم الحصول على الوسيط بموجب المعادلة الآتية:

$$\frac{\text{الوسيط} - ١٣}{١٤-١٣} = \frac{١٣-٩}{١٧-٩}$$

$$\frac{\text{الوسيط} - ١٣}{١٤-١٣} = \frac{٤}{٨}$$

8 الوسيط - 104 = 56 - 52

8 الوسيط - 104 = 4

= 104 + 4 = 108

الوسيط = 13.5

(جـ) حساب الوسيط من جدول الفئات التكرارية:

إذا كان لدينا جدول تكراري (يتضمن فئات بتكراراتها)، فإن حساب الوسيط يتطلب تحديد الفئة الوسيطية (أي الفئة التي يقع فيها الوسيط)، ويتم حساب الوسيط باستخدام التوزيع التكراري المتجمع الصاعد أو الهابط.

1) باستخدام التوزيع التكراري المتجمع <u>الصاعد</u>: في هذه الحالة يكون الوسيط بموجب المعادلة الآتية:

$$\text{الوسيط} = \text{ل} + \left(\frac{\text{ت} - \text{ك}_{\text{م}}}{\text{ك}} \times \text{ف} \right)$$

حيث إن:

ل : الحد الأدنى للفئة الوسيطية

ت: ترتيب الوسيط

ك م: التكرار المتجمع السابق للفئة الوسيطية

ك: تكرار الفئة الوسيطة

ف: طول الفئة الوسيطية

وبناء على ذلك، لابد من الحصول على كل المعطيات السابقة من التوزيع التكراري المتجمع الصاعد.

مثال:

الجدول الآتي يوضح توزيع عدد من حلقات برامج إذاعية حسب المدة الزمنية بالدقيقة:

الفئات	التكرارات
٣٠-٤٠	٢٧
٤٠-٥٠	١٢
٥٠-٦٠	٣٢
٦٠-٧٠	٢٣
٧٠-٨٠	٢٢
المجموع	١١٦

للحصول على الوسيط من هذا الجدول باستخدام التوزيع التكراري المتجمع الصاعد، فإن نقطة البدء هي إيجاد هذا التوزيع، ثم تحديد ترتيب الوسيط، وكذلك الفئة الوسيطية:

الفئات	التكرارات	التكرار المتجمع الصاعد
٣٠-٤٠	٢٧	٢٧
٤٠-٥٠	١٢	٣٩
٥٠-٦٠	٣٢	٧١
٦٠-٧٠	٢٣	٩٤
٧٠-٨٠	٢٢	١١٦

وإذا كان مجموع التكرارات هـو (١١٦)، فـإن ترتيـب الوسـيط = ١١٦ ÷ ٢ = ٥٨ ومعنـى ذلـك أن الوسـيط يقـع بـين القيمـة (٣٩) والقيمـة (٧١) في التوزيـع التكـراري المتجمـع الـصاعد، وتكـون الفئـة الوسيطية هي تلك الفئة التـي يقـع فيهـا الوسـيط، وهـي في مثالنـا الفئـة (مـن ٥٠ إلى ٦٠)، ويوضـح الجدول أن الحد الأدنى للفئة الوسيطية هو (٥٠) أما التكرار المتجمع للفئة السابقة عليهـا (التـي قبـل الفئة الوسيطية)، فهو (٣٩)، كما أن تكـرار الفئـة الوسـيطية يـساوي (٣٢) وطولهـا (١٠)، وبـالعودة إلى معادلة الوسيط وهي:

$$م = ل + \left(\frac{ت - ك م}{ك} \times ف \right)$$

وباستخدام المعطيات التي حصلنا عليها، والتعويض في هذه المعادلة، فإن:

$$م = ٥٠ + \left(\frac{58 - 39}{32} \times 10 \right) = ٥٦$$

أي أن الوسيط هو ٥٦ تقريبًا

2) باستخدام التكرار المتجمع الهابط: في هذه الحالة تستخدم المعادلة الـسابق ذكرهـا، ولكـن مـع تغيير طفيف كالآتي:

$$الوسيط = ل - \left(\frac{ت - ك مل}{ك} \times ف \right)$$

حيث: ل: الحد الأعلى للفئة الوسيطية

ت: ترتيب الوسيط

ك م ل: التكرار المتجمع للفئة التي تلي الفئة الوسيطة

ك: تكرار الفئة الوسيطية

ف: طول الفئة الوسيطية

وبتكوين التوزيع التكراري المتجمع الهابط يكون لدينا التوزيع التالى:

الفئات	التكرارات	التكرار المتجمع الهابط
٣٠-٤٠	٢٧	١١٦
٤٠-٥٠	١٢	٨٩
٥٠-٦٠	٣٢	٧٧
٦٠-٧٠	٢٣	٤٥
٧٠-٨٠	٢٢	٢٢
المجموع	١١٦	

وقد سبقت الإشارة إلى أن ترتيب الوسيط هو (٥٨) كما أن الجدول يوضح أن الحد الأعلى للفئة الوسيطية هو (٦٠)، ويوضح أيضًا أن التكرار المتجمع الهابط للفئة التي تلي الفئة الوسيطية هو (٤٥)، كما أن تكرار الفئة الوسيطية هو (٣٢) وطولها (١٠)، وباستخدام هذه المعطيات والتعويض في المعادلة المذكورة، فإن الوسيط يساوي:

$$٦٠ - \left(10 \times \frac{45 - 58}{32}\right)$$

٦٠- ٤ = ٥٦ تقريبًا

أي أن الوسيط يساوي ٥٦

مثال آخر:

في دراسة ارتكزت على تحليل محتوى نشرات الأخبار في إذاعة الرياض بالمملكة العربيـة الـسعودية، خلص الباحث إلى نتائج هامة من بينها أن عينة النشرات كانت تتوزع حـسب المـدة الزمنيـة بالدقيقـة وفق ما يوضحه الجدول الآتي:

الفئات	التكرارات
٢٠-٢٤	٢٨
٢٥-٢٩	٣٦
٣٠-٣٤	٣١
٣٥-٣٩	٢٦
٤٠-٤٤	١٥
المجموع	**١٣٦**

إن الحصول على الوسيط باستخدام التكرار المتجمع الـصاعد (ت م ص) وكـذلك باسـتخدام التكـرار المتجمع الهابط (ت م هـ)، يبدأ بتكوين كلا التوزيعين وفق ما يوضحه الجدول الآتي:

الفئات	ك	ت م ص	ت م هـ
٢٠-٢٤	٢٨	٢٨	١٣٦
٢٥-٢٩	٣٦	٦٤	١٠٨
٣٠-٣٤	٣١	٩٥	٧٢
٣٥-٣٩	٢٦	١٢١	٤١
٤٠-٤٤	١٥	١٣٦	١٥
	١٣٦		

بما أن مجموع التكرارات هو (١٣٦)، فتكون رتبة الوسيط تساوي ١٣٦ ÷ ٢ = ٦٨

فإذا استخدمنا التوزيع التكراري المتجمع الصاعد (ت م ص)، فإن رتبة الوسيط تقع بعد القيمة (٦٤) في هذا التوزيع؛ أي أن الوسيط يقع في الفئة (٣٠-٣٤) وهذه هي الفئة الوسيطية، كما نتبين من الجدول أن التكرار المتجمع الصاعد للفئة التي تسبق الفئة الوسيطية هو (٦٤)، أما طول الفئة الوسيطية، فإنه يساوي (٥)، وحدها الأدنى (٣٠) وتكرارها الأصلي ٣١ وهكذا يتم حساب الوسيط بتطبيق المعادلة:

$$\text{م} = \text{ل} + \left(\frac{\text{ت} - \text{ك م}}{\text{ك}} \times \text{ف} \right)$$

وبالتالي يكون الوسيط:

$$\text{م} = ٣٠ + \left(5 \times \frac{64-68}{31} \right)$$

= ٣٠ + ٠.٧ = ٣٠.٦٥

أما إذا استخدمنا التوزيع التكراري المتجمع الهابط، فإننا نطبق المعادلة السابقة بالصيغة الآتية:

$$\text{الوسيط} = \text{ل} - \left(\frac{\text{ت} - \text{ك مل}}{\text{ك}} \times \text{ف} \right)$$

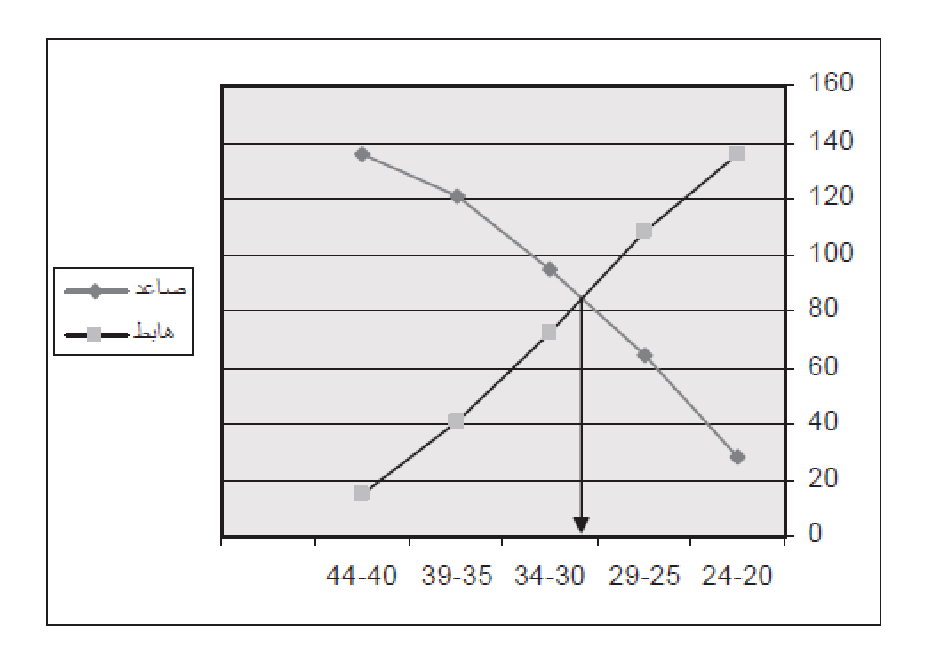
وبما أن ترتيب الوسيط = ٦٨ وبما أن الفئة الوسيطية (٣٠-٣٥)، فإن طولها يساوي (٥) وتكرارها الأصلي ٣١ وحدها الأعلى ٣٤ كما أن التكرار المتجمع الهابط للفئة التي تلي الفئة الوسيطية هو (٤١). باستخدام هذه المعطيات. وبالتعويض في المعادلة السابقة يكون الوسيط يساوي:

$$٣٥ - \left(5 \times \frac{41-68}{31} \right)$$

= ٣٥ – ٤.٣ = ٣٠.٦٥

ويلاحظ أن قيمة الوسيط واحدة في الحالتين، سواء عند استخدام التوزيع التكراري المتجمع الصاعد (ت. م. ص.) أو عند استخدام التوزيع التكراري المتجمع الهابط (ت. م. هـ).

ويمكن الحصول على الوسيط بيانيًا من خلال رسم المنحنى التكراري المتجمع الصاعد والمنحنى التكراري المتجمع الهابط، وعند نقطة تقاطع المنحنيين نسقط عمودًا رأسيًا مستقيمًا على المحور الأفقي، وتكون قيمة الوسيط هي القيمة التي يلتقي فيها العمود مع المحور الأفقي، فيما يوضحه الرسم الآتي:



ومن خصائص الوسيط أنه لا يتأثر بالقيم المتطرفة، فهو يتم حسابه من البيانات المبوبة أو التي تم ترتيبها بصرف النظر عن وجود تلك القيم من عدمه، كما أن الوسيط لا يتأثر بالفئات المفتوحة، ويفيد الوسيط كثيرًا في رصد متوسط الظاهرة عبر فترة زمنية معينة، كما أن الوسيط لا يعتمد على مراكز الفئات، بل يعتمد على التكرارات فقط، ولذلك فهو

يستخدم - دون غيره- في الفئات المفتوحة، كما يصلح استخدام الوسيط عندما يكون توزيع التكرارات ملتويّا دلالة على تطرف البيانات وانحيازها في اتجاه معين.

ثالثًا: المنوال

المنوال (Mode) هو القيمة الأكثر شيوعًا، أو الأكثر تكرارًا، افترض أن لدينا عينة من ٥٠٠ شخص، وكانت هذه العينة تتوزع حسب ساعات المشاهدة اليومية للتليفزيون كالآتي:

- أقل من ساعة ٦٨
- من ساعة إلى ساعتين ٣٩٧
- أكثر من ساعتين ٣٥

هنا تكون الفئة الثانية (من ساعة إلى ساعتين) هي الفئة المنوالية؛ لأنها تكررت أكثر من غيرها؛ إذ إنها استحوذت على ٣٩٧ تكرارًا؛ أي ما يعادل ٨٠% تقريبًا من العينة، ويمكن حساب المنوال من البيانات الخام مباشرة أو من البيانات المبوبة في جداول تكرارية.

(أ) حساب المنوال من البيانات الخام:

يتم حساب المنوال من البيانات الخام مباشرة من خلال معرفة عدد المرات التي تكرر فيها كل رقم، لنفرض أن البيانات الآتية تمثل درجات خمسة وعشرين مبحوثًا على مقياس الاستخدامات والإشباعات:

١٥، ١٢، ١٣، ١٢، ١٧، ١٧، ١٧، ١٧، ٢٠ ، ١٧، ٢٣، ١٥، ١٧،

١٠، ١٣، ١٢، ١٧، ١٨، ١٧، ٢٢، ١٤، ١٥، ١٧، ١٧، ١٧، ١٢

إن حساب المنوال لمثل هذه البيانات يكون من خلال معرفة تكرارات كل رقم، ومن البيانات السابقة يتضح أن القيمة (١٧) تكررت أكثر من القيم الأخرى؛ أي أن المنوال هو

(١٧)؛ ولتسهيل حساب المنوال يمكن ترتيب الأرقام تصاعديّا أو تنازليّا.

وقد يجد الباحث أن لديه بيانات لها أكثر من منوال، كأن تكون قيمتان أو أكثر لها العدد نفسه من التكرارات وبما يزيد عن تكرارات القيم الأخرى، مثال: ٢، ٣، ٣، ٣، ٣، ٤، ٤، ٤، ٥، ٥، ٦، ٦، ٦، ٦، ٦ فهذه البيانات لها منوالان هما (٣)، (٦)، فالقيمة (٣) تكررت أربع مرات، وكذلك القيمة (٦) تكررت أيضًا أربع مرات، كما يلاحظ من هذا المثال أن القيم المنوالية غير متجاورة، علمًا بأنه في حالة وجود قيم متجاورة ومتساوية في التكرارات بما يزيد على تكرارات القيم الأخرى يتم حساب المنوال بجمع القيم ذات التكرارات الأعلى وقسمة الناتج على ٢ مثال ذلك: ٨، ٨، ٨، ٨، ٨، ٩، ٩، ٩، ٩، ٩، ١٠، ١٠، ١٠، ١١، ١١ فالرقم (٨) تكرر خمس مرات، وهو يجاور الرقم (٩) الذي تكرر هو الآخر خمس مرات، وهناك يكون المنوال ٨+ ٩ ÷ ٢ = ٨.5

في الوقت نفسه قد يجد الباحث أن لديه بيانات ليس لها منوال؛ حيث لا توجد قيمة تكررت أكثر من غيرها، مثال ذلك: ١٨، ٢٢، ٢٣، ٣٧، ٤٤، ٤٥ ففي هذه البيانات لا توجد قيمة تكررت أكثر من غيرها، وبالتالي ليس لها منوال. وينطبق ذلك أيضًا على البيانات التي تتكرر فيها القيم أكثر من مرة بدرجة متساوية، مثال ذلك: ٦٨، ٦٨، ٦٨، ٧٣، ٧٣، ٧٣، ٩٤، ٩٤، ٩٤ فهذه البيانات لا يوجد لها منوال لأنها تكررت بأعداد متساوية (فالعدد ٦٨ تكرر ثلاث مرات وكذلك العدد ٧٣ والعدد ٩٤)

(ب) حساب المنوال من بيانات مبوبة:

كثيرًا ما يكون المطلوب إيجاد المنوال من بيانات مبوبة في جدول تكراري يتضمن الفئات والتكرارات الخاصة بكل فئة، وهناك حالات متنوعة لذلك، وبالتالي هناك عدة قوانين للحصول على المنوال، وسوف نقدم مثالًا واحدًا لتوضيح طريقة الحصول على المنوال من بيانات مبوبة، وهي طريقة الفروق لكارل بيرسون. ففي دراسة عن المعلومات الطبية التي اكتسبها الجمهور من وسائل الاتصال تم تطبيق اختبار مقنن، وكانت العينة تتوزع حسب فئات الدرجات على هذا الاختبار وفق ما يوضحه الجدول الآتي:

فئات الدرجات	التكرارات
١- ٢٠	١١
٢٠- ٤٠	١١٤
٤٠- ٦٠	١٨٩
٦٠- ٨٠	٣٤
٨٠- ١٠٠	٥٢
المجموع	**٤٠٠**

إن مجرد النظر في هذا الجدول يكشف عن أن الفئة المنوالية هي الفئة الثالثة (٤٠- ٦٠) فهي الفئة التي يوجد فيها أكبر عدد من التكرارات (١٨٩ تكرارًا)، وبموجب معرفة الفئة المنوالية في الجدول التكراري يمكن إيجاد المنوال بواسطة القانون:

$$\text{المنوال} = \text{الحد الأدنى للفئة المنوالية} + \frac{\text{الفرق الأول}}{\text{الفرق الأول} + \text{الفرق الثاني}} \times \text{طول الفئة المنوالية}$$

إن الفرق الأول، يعني الفرق بين تكرارات الفئة المنوالية وتكرارات الفئة السابقة عليها؛ أي ١٨٩- ١١٤ = ٧٥

أما الفرق الثاني، فهو الفرق بين تكرارات الفئة المنوالية وتكرارات الفئة التالية لها مباشرة؛ أي ١٨٩- ٣٤ = ١٥٥

أي أن قيمة الفرق الأول تبلغ ٧٥ أما قيمة الفرق الثاني، فهي ١٥٥ ومجموعهما هو ٢٣٠ أما طول الفئة المنوالية، فهو ٢٠ كما أن الحد الأدنى للفئة المنوالية هو ٤٠

وبتطبيق القانون السابق يكون المنوال يساوي:

$$\text{المنوال} = ٤٠ + \frac{75}{230} \times ٢٠ = ٥٢.٥$$

أي أن المنوال في هذا المثال هو ٥٢.٥ وكما سبقت الإشارة، فإن هذه الطريقة تعرف بطريقة الفروق لكارل بيرسون، وهناك صور متعددة لهذه الطريقة، وإن كانت الطريقة المذكورة هي الأكثر شيوعًا في البحوث.

كما يمكن الحصول على المنوال بمعلومية الوسط الحسابي والوسيط؛ حيث يستخدم القانون المعروف:

المنوال = ٣ (الوسيط)- ٢ (الوسط الحسابي)

فإذا كان الوسيط ٣٦ والوسط الحسابي ٣٣، فإن المنوال يكون:

٣ (٣٦) - ٢(٣٣) = ١٠٨ – ٦٦ = ٤٢

خصائص المنوال:

يمتاز المنوال بأنه من مقاييس النزعة المركزية سهلة الفهم والحساب، ولا يتأثر بالقيم المتطرفة إذا كانت بعيدة عن الفئة المنوالية، ويفيد استخدامه في رصد التكرارات الشائعة، الأمر الذي يجعله مفيدًا في الكثير من التطبيقات، مثل معرفة الوسائل الإعلامية والبرامج والمواد التي يفضلها أو لا يفضلها الجمهور. ويتطلب المنوال تنظيم البيانات في صورة توزيع تكراري، ولا يعتمد المنوال في حسابه على كل المفردات وإنما يقتصر على الفئة المنوالية. غير أن المنوال يفقد معناه عندما تكون المفردات كثيرة العدد، قريبة من بعضها البعض.

العلاقة بين مقاييس النزعة المركزية:

هناك علاقة وثيقة بين الوسط الحسابي والوسيط والمنوال، وتتأثر هذه العلاقة بشكل التوزيع التكراري (معتدل، ملتوٍ ناحية اليمين أو ناحية اليسار)؛ إذ إن شكل التوزيع يؤثر في قيمة هذه المقاييس، ففي التوزيعات التكرارية المتماثلة تتفق قيم الوسط الحسابي والوسيط والمنوال، أما في التوزيعات التكرارية غير المتماثلة، فإن العلاقة بين هذه المقاييس تكون بموجب المعادلة:

المنوال= ٣ (الوسيط)- ٢ (الوسط الحسابي)

فإذا كان الوسيط ٣٦ والوسط الحسابي ٣٣، فإن المنوال يكون:

٣ (٣٦) - ٢(٣٣) = ٤٢

كما يمكن إيجاد قيمة الوسط الحسابي، بمعلومية الوسيط والمنوال بموجب المعادلة:

الوسط الحسابي = ٣× الوسيط-٢× المنوال

فإذا كان الوسيط ١٨ والمنوال ١٦، فإن الوسط الحسابي يكون:

٣×١٨ – ٢×١٦ =٢٢

غير أن العلاقة بين مقاييس النزعة المركزية تتخذ أشكالًا أخرى حسب طبيعة التوزيع ودرجة عدم تماثله، أو ميل التوزيع إلى اليمين أو إلى اليسار.

المطلب الثاني: مقاييس التشتت

تدلنا مقاييس النزعة المركزية على تجمع البيانات، والقيم المتوسطة لها، غير أن ذلك لا يعتبر كافيًا لوصف البيانات بدقة، فقد يتساوى الوسط الحسابي مثلًا لمجموعتين من البيانات، رغم وجود فروق كبيرة في القيم الخاصة بكل مجموعة:

فالوسط الحسابي للقيم ٦، ٩، ١٢ هو ٩

والوسط الحسابي للقيم ١، ٢، ٢٤ هو أيضًا ٩

أي أن قيمة الوسط الحسابي هي نفسها في المجموعتين رغم وجود اختلافات واضحة في أعداد كل مجموعة، فبينما تتكون المجموعة الأولى من الأعداد ٦، ٩، ١٢ نجد أن المجموعة الثانية تتكون من أعداد مختلفة تمامًا وهي ١، ٢، ٢٤ من هنا، فإن مقاييس النزعة المركزية – وإن كانت تعطي وصفًا واضحًا للبيانات – إلا أن هذا الوصف يظل محددًا

بتلك النزعة، ولا توضح تلك المقاييس ما هو أبعد من ذلك، وبالتالي يتعين استخدام مقاييس إضافية للوصول إلى وصف دقيق للبيانات، هذا الوصف تتيحه مقاييس التشتت، ومن أمثلتها: المدى الكلي، التباين والانحراف المعياري، الانحراف الرُّبيعي، متوسط الانحرافات، معامل الالتواء، معامل الاختلاف، الدرجات المعيارية:

أولًا: المدى الكلي

المدى الكلي (Total Range) هو الفرق بين أصغر قيمة وأكبر قيمة مع إضافة واحد صحيح لهذا الفرق.

مثـال:

فيما يلي بيان بعدد حلقات البرامج الإذاعية التي أنتجتها بعض الشركات المتخصصة في إنتاج هذه النوعية من البرامج:

الشركة المنتجة	عدد حلقات البرامج الإذاعية
الطارق للإنتاج الفني	٧٨
بين القصرين	٦٥
النيل للإعلام	٤٤
الوعي	٨٩
الشعلة	٦٧
الصفا والمروة	٨٩
العروة الوثقى	٧٣
مؤسسة المنار	١٤٢
الخطوة	٥٦
الضوء الشارد	٥٨
المجموع	**٧٦١**

ولمعرفة المدى الكلي، فإننا نلاحظ من هذه البيانات أن أصغر قيمة هي ٤٤ (وهي عدد حلقات البرامج الإذاعية التي أنتجتها شركة النيل للإعلام)، أما أكبر قيمة فهي ١٤٢ (وهي عدد حلقات البرامج الإذاعية التي أنتجتها مؤسسة المنار)، ويكون المدى الكلي هو:

(١٤٢- ٤٤) + ١ = ٩٩

أي أن المدى الكلي للبيانات التي يتضمنها الجدول هو ٩٩، ويتضح من ذلك أن المدى الكلي لا يعتمد على كل القيم العددية، وإنما يعتمد فقط على قيمتين: القيمة الصغرى والقيمة الكبرى وهذه نقطة الضعف الأساسية في المدى الكلي كأحد مقاييس التشتت، الأمر الذي يجعله غير دقيق.

ويمكن استخدام المدى الكلي للمقارنة بين تشتت التوزيعات التكرارية المختلفة شريطة أن تكون تلك التوزيعات متساوية من حيث العدد، فمن الممكن استخدام المدى الكلي لمقارنة التشتت لعينة تضم ٤٠٠ مفردة، وعينة أخرى تضم العدد نفسه من المفردات (٤٠٠ مفردة)، لكن لا يصلح استخدام المدى الكلي لمقارنة التشتت بين تكرارات عينتين مختلفتين من حيث الحجم، خاصة إذا كان الفارق بينهما كبيرًا (كأن تكون العينة الأولى ١٠٠ والعينة الثانية ٥٠٠ مثلًا).

أما إذا كانت البيانات موزعة على فئات في جدول تكراري، فإن المدى الكلي = الحد الأعلى للفئة الأخيرة منقوصًا منه الحد الأدنى للفئة الأولى، كما يتضح من المثال المبين في الجدول التالي:

الفئات	ك
٥-٩	٨
١٠- ١٤	١٢
١٥- ١٩	٧
٢٠- ٢٤	٨
المجموع	**٣٥**

فالحد الأعلى للفئة الأخيرة هو (٢٤) والحد الأدنى للفئة الأولى هو (٥)، وعلى ذلك يكون المدى الكلي = ٢٤- ٥ = ١٩

غير أن المدى الكلي - كمقياس للتشتت - غير شائع الاستخدام في البحوث العلمية؛ لأنه يعتمد فقط على القيمتين الأصغر والأكبر.

ثانيًا: التباين والانحراف المعياري

التباين (Variance) هو متوسط مجمل مربعات الانحراف عن الوسط الحسابي، أما الانحراف المعياري (Standard Deviation)، فهو الجذر التربيعي لمتوسط مجموع مربعات الانحرافات عن الوسط الحسابي؛ أي أن التباين هو مربع الانحراف المعياري (وبالتالي فإن الانحراف المعياري هو الجذر التربيعي للتباين).

ويعتبر التباين والانحراف المعياري من أهم مقاييس التشتت، ويعتمدان على الوسط الحسابي بشكل أساسي، فإذا افترضنا أن هناك مقياسًا لقياس استخدام التليفزيون، وكان متوسط درجة المفحوصين على هذا المقياس هو (٣٦.٣٢) درجة، فإن هناك مفحوصين حصلوا على درجة تقل عن هذا المتوسط،وهناك أيضًا مفحوصون آخرون حصلوا على درجة تزيد على المتوسط، هذه الفروق (الانحراف عن الوسط الحسابي للعينة) نعتمد عليها في حساب التباين والانحراف المعياري للعينة ككل.

ومن الناحية الإحصائية، فإن مجموع الانحرافات عن الوسط الحسابي يساوي الصفر؛ لأن مجموع الفروق السالبة يتساوى مع مجموع الفروق الموجبة، ومن هنا يتم (تربيع) جميع الفروق حتى لا يكون مجموعها صفرًا، فلنتأمل المثال الآتي:

لدينا بيان بعدد الساعات التي يقضيها خمسة طلاب في قراءة الصحف أسبوعيا، وهذه الساعات هي: ٢، ٣، ٤، ٥، ٦ فالطالب الأول يقضي ساعتين، أما الطالب الثاني فيقضي ثلاث ساعات... إلخ. إن الوسط الحسابي لهذه الأعداد = ٢+٣+٤+٥+٦÷ ٥=٤

أي أن متوسط وقت قراءة الصحف أسبوعيًا هو ٤ ساعات، ويتم حساب انحراف كل قيمة عن الوسط الحسابي، وذلك على النحو الآتي:

- انحراف القيمة الأولى = ٢-٤=-٢
- انحراف القيمة الثانية = ٣-٤=-١
- انحراف القيمة الثالثة = ٤-٤=صفر
- انحراف القيمة الرابعة ٥-٤-= ١
- انحراف القيمة الخامسة ٦-٤= ٢

فإذا حسبنا مجمل ناتج تلك الانحرافات، نجد أنه يساوي الصفر:

-٢-١+ صفر +١+٢ = صفر

من هنا جاءت رؤية «كارل بيرسون» بتربيع الفروق عن الوسط؛ وذلك للتخلص من العلامات السالبة، وبتربيع الفروق عن الوسط، نجد أن القيم السابقة تصبح:

- مربع فروق القيمة الأولى = -٢×-٢ = ٤
- مربع فروق القيمة الثانية = -١×-١= ١
- مربع فروق القيمة الثالثة = صفر× صفر= صفر
- مربع فروق القيمة الرابعة = ١ × ١ = ١
- مربع فروق القيمة الخامسة= ٢×٢ = ٤

وبجمع الناتج نجد أنه يساوي ٤+١+صفر +١+٤=١٠

ونظرًا لأن عدد الحالات هو خمسة، فإن متوسط هذا الناتج هو ١٠ ÷ ٥ = ٢ وبذلك يكون متوسط مجموع مربعات الانحرافات يساوي ٢ وهذه القيمة هي التباين، ويكون الانحراف المعياري هو جذر تلك القيمة، بمعنى أن الانحراف المعياري = $\overline{2}$ b = ١.٤١ أي أن التباين هو مربع الانحراف المعياري:

التباين = $ع^{٢}$

فإذا كان الانحراف المعياري لعدد ساعات قراءة الصحف هو ١.٤١ كما سبقت الإشارة، فإن التباين يساوي $(١.٤١)^{٢}$ ؛ أي ١.٤١×١.٤١= ٢ تقريبًا

بموجب ذلك يمكن معرفة قيمة التباين بمعلومية قيمة الانحراف المعياري، كما يمكن معرفة قيمة الانحراف المعياري بمعلومية قيمة التباين، فإذا افترضنا أن دراسة أجريت على عينة من قراء الصحف، وكان الانحراف المعياري للوقت المنقضي في قراءة الصحف اليومية ٢.٧٢، فإن قيمة التباين تساوي($ع^{٢}$) أو مربع الانحراف المعياري، وبما أن قيمة الانحراف المعياري ٢.٧٢، فإن التباين هو $(٢.٧٢)^{٢}$ أي ٢.٧٢ ×٢.٧٢ = ٧.٤

حساب التباين والانحراف المعياري من البيانات المبوبة:

إذا كان لدى الباحث بيانات ينتظمها جدول تكراري (فئات وتكرارات)، فإن حساب التباين (وبالتالي الانحراف المعياري) يكون بموجب المعادلة:

$$ع^{٢} = \frac{\text{مجـ}ص^{2}ك - \frac{(\text{مجـ}ص\,ك)^{2}}{\text{مجـ}\,ك}}{\text{مجـ}ك - 1}$$

حيث إن:

$ع^{٢}$: التباين

ص : منتصفات الفئات

$ص^{٢}$: مربعات منتصفات الفئات

ك : عدد المفردات

وبالحصول على التباين (ع^{٢}) نحصل على الانحراف المعياري باعتباره الجذر التربيعي للتباين.

مثال: الجدول الآتي يوضح توزيع عينة من القائمين بالاتصال في إحدى إذاعات (FM) حسب مدة الخبرة في العمل الإعلامي بالسنوات:

المدة الزمنية	ك
٢-٤	٥
٥-٧	٧
٨- ١٠	٣
المجموع	**١٥**

نبدأ بحساب منتصفات الفئات، وذلك بجمع بداية ونهاية كل فئة ثم قسمة الناتج على ٢ (فمنتصف الفئة الأولى مثلًا = ٢+٤÷ ٢ = ٣)، وبتربيع منتصفات الفئات والضرب في التكرارات نحصل على الجدول الآتي:

المدة الزمنية	ك	ص	ص^{٢}	ص × ك	ص^{٢} × ك
٢-٤	٥	٣	٩	١٥	٤٥
٥-٧	٧	٦	٣٦	٤٢	٢٥٢
٨- ١٠	٣	٩	٨١	٢٧	٢٤٣
المجموع	**١٥**			**٨٤**	**٥٤٠**

بموجب ذلك يمكن الحصول على التباين والانحراف المعياري، وذلك باستخدام المعادلة المشار إليها وهي:

$$ع^{٢} = \frac{\text{مجص}^2\text{ك} - \frac{(\text{مجـ ص ك})^2}{\text{مجـك}}}{\text{مج ك} - 1}$$

وبالتعويض في هذه المعادلة يكون:

$$ع^٢ = \frac{540 - \frac{(84)^2}{15}}{15 - 1}$$

$$ع^٢ = \frac{540 - 470.4}{14} = ٤.٩٧$$

أي أن التباين = ٤.٩٧

وبالتالي يكون الانحراف المعياري = $b\sqrt{4.97}$ = ٢.٢٣

خصائص التباين والانحراف المعياري:

إن التباين يقيس التشتت من خلال مربعات الانحرافات عن الوسط الحسابي؛ أي أنه يعتمد في حسابه على قيم غير أصلية ، ويأخذ في اعتباره جميع المفردات، ولا تتأثر قيمة التباين بإضافة أو طرح أي قيمة ثابتة على جميع القيم المحسوب منها قيمة التباين، لكن التباين والانحراف المعياري يتأثران بالقيم المتطرفة (لاعتمادهما على مربع انحراف هذه القيم عن الوسط الحسابي)، كما يتأثران بقيمة الوسط الحسابي، الذي بدوره يتأثر بالقيم المتطرفة.

وإذا كان الانحراف المعياري يعتمد على التباين، فإن الانحراف المعياري من أهم مقاييس التشتت، كما أنه يرتبط ارتباطًا وثيقًا بمقاييس إحصائية هامة مثل معامل الالتواء والارتباط والدرجات المعيارية والدلالة الإحصائية... إلخ. وعلى وجه التحديد، فإن الانحراف المعياري يتصف بالخواص الآتية:

- إن قيمة الانحراف المعياري لا تتأثر بحذف أو إضافة عدد ثابت لكل قيمة من قيم التوزيع التكراري، ولتوضيح ذلك نأخذ المثال التالي:

انحراف العدد ٨ عن العدد ١٢ يساوي ٤ ؛ أي ١٢- ٨= ٤

فإذا أضفنا عـددًا ثابتًا (وليكن ٣) إلى العدد ٨، فإن النـاتج يكـون (١١)، وإذا أضـفنا العـدد (٣) إلى العدد (١٢) يكون الناتج (١٥)، وبحساب الانحراف بعد إضافة العدد الثابـت نجـد أنـه = ١٥- ١١ = ٤ وهي نفس قيمة انحراف العدد ٨ عن العدد ١٢ وهكذا. المنطق نفسه إذا أنقصنا عددًا ثابتًا (وليكن ٤) من العدد ٨ يكون الناتج ٤ وإذا أنقصنا العدد ٤ من العدد ١٢، يكون الناتج ٨، بحساب الانحراف بعـد إنقاص العدد الثابت نجد أنه = ٨- ٤ = ٤ وهي نفس قيمة انحراف العدد ٨ عن العدد ١٢، هكذا نتبين أن الانحراف لا يتأثر بإضافة أو حذف عدد ثابت لكل قيمة من قيم التوزيع التكراري.

- إن القيمة الجبرية للانحراف المعياري يمكـن أن تكـون سـالبة أو موجبـة، وبالتـالي فـإن الانحـراف المعياري يقيس التشتت بالانحرافـات التـي عـلى نـاحيتي الوسـط بالنـاقص أو بالزائـد (الناحيـة اليسرى والناحية اليمنى) للتوزيع الاعتدالي.

- إن الانحراف المعياري يقسم قاعدة منحنـى التوزيـع التكـراري إلى أقـسام متـساوية، الأمـر الـذي يجعله أكثر دقة ووضـوحًا في قيـاس التـشتت، وعنـدما يتـضمن التوزيـع التكـراري عـددًا كبـيرًا، ويكون توزيعًا اعتداليًا (Normal Distribution)، فإن الانحراف المعياري يقـسم المـدى الكـلي إلى ستة أقسام متساوية، وعلى هذا الأساس فإن هناك علاقة بين المدى الكلي والانحراف المعياري؛ إذ إن:

الانحراف المعياري = المدى الكلي ÷ ٦ (تقريبًا)، ففي مثالنا السابق عن المدى الكـلي لعـدد حلقـات البرامج الإذاعية التي أنتجتها الشركات الإعلامية، كانت قيمة هذا المدى هي ٩٩ ومن هذه القيمـة فـإن الانحراف المعياري يساوي:

٩٩÷ ٦ = ١٦.٥

مع التنبيه بأن ذلك يكون في التوزيعات التكرارية الاعتدالية (وليس التوزيعات الملتوية).

ثالثًا: الانحراف الرُّبيعي

يتم حساب الانحراف الرُّبيعي (نصف المدى الرُّبيعي) من خلال الرُّبيع الأول والرُّبيع الثالث. فالرُّبيع الأول هو تلك القيمة التي تقل عنها ٢٥% من القيم. أما الرُّبيع الثالث فهو تلك القيمة التي تقل عنها ٧٥% من القيم (لاحظ أن الرُّبيع الثاني هو الوسيط)، ولما كان الانحراف الرُّبيعي يعتمد على الرُّبيع الأول والرُّبيع الثالث، فإنه لا بد من الحصول عليهما أولًا، وذلك من خلال تحديد ترتيب كل منهما بين البيانات محل الدراسة، ثم حساب قيمة كل منهما، ومن خلال ترتيب جميع القيم تصاعديًا (أو تنازليًا) يمكن بسهولة تحديد ترتيب الرُّبيع الأول وترتيب الرُّبيع الثالث:

ففيما يخص ترتيب الرُّبيع الأول يكون من خلال عدد القيم ÷ ٤

أي $\frac{ن}{4}$

وفيما يخص ترتيب الرُّبيع الثالث يكون من خلال عدد القيم × ٣ ÷ ٤

أي $\frac{ن \times 3}{4}$

حيث (ن) ترمز إلى عدد المفردات، ومتى حددنا ترتيب الرُّبيع الأول، وترتيب الرُّبيع الثالث- بمعنى موقع كل منهما من ترتيب المفردات - يصبح من السهل الحصول على قيمة كل منهما، ويكون الانحراف الرُّبيعي ناتج قيمة الرُّبيع الثالث مطروحًا منها قيمة الرُّبيع الأول والقسمة على ٢ أي أن:

$$\text{الانحراف الرُّبيعي} = \frac{\text{الرُّبيع الثالث} - \text{الرُّبيع الأول}}{٢}$$

مثال:

فيما يلي بيان بالإعلانات الدينية المنشورة في عشر صحف يومية خلال شهر رمضان:

٢٠ – ١١ – ١٥ – ١٣ – ١٢ – ١٨ – ١٩ – ١٦ – ١٠ – ١٤

بترتيب هذه القيم تصاعديًا يكون:

١٠ – ١١ – ١٢ – ١٣ – ١٤ – ١٥ – ١٦ – ١٨ – ١٩ – ٢٠

ويلاحظ أن هذه القيم عددها (١٠)، أي أن قيمة (ن) = ١٠

ويكون ترتيب الرُّبيع الأول = $\frac{ن}{4}$ = $\frac{10}{4}$ = ٢.٥

وبهذا، فإن الرُّبيع الأول ترتيبه (٢.٥)؛ أي أنه يقع بين المفردة الثانية (وقيمتها ١١) والمفردة الثالثة (وقيمتها ١٢)، فتكون قيمة الرُّبيع الأول تساوي:

$$\frac{12+11}{2} = ١١.٥$$

ترتيب الرُّبيع الثالث = $\frac{3\times ن}{4}$ = $\frac{3\times10}{4}$ = ٧.٥

أي أن ترتيب الرُّبيع الثالث هو (٧.٥)؛ أي أنه يقع بين المفردة السابعة (وقيمتها ١٦) والمفردة الثامنة (وقيمتها ١٨)، وبالتالي تكون قيمة الرُّبيع الثالث تساوي:

$$\frac{18+16}{2} = ١٧$$

هكذا يتضح أن قيمة الرُّبيع الأول = ١١.٥ كما أن قيمة الرُّبيع الثالث = ١٧ وبالتالي فإن:

الانحراف الرُّبيعي = $\frac{17-11.5}{2}$ = ٢.٧٥

رابعًا:متوسط الانحرافات

وهو متوسط انحرافات القيم عن وسطها الحسابي، مع تجاهل الإشارات السالبة (علامة الناقص)، وهذه نقطة ضعف أساسية في هذا الأسلوب كمقياس للتشتت. كمثال توضيحي، نفرض أن لدينا القيم الأربع التالية: (١٦)، (٢٤)، (٢٠) ، (٢٨). إن مجموع هذه القيم هو (٨٨)، ووسطها الحسابي = ٨٨÷٤= ٢٢

ويكون متوسط الانحرافات يساوي:

(١٦-٢٢) + (٢٤-٢٢) + (٢٠-٢٢) + (٢٨-٢٢) ÷٤

= (- ٦) + (٢) + (-٢) + (٦) ÷ ٤

وبتجاهل الإشارات السالبة، فإن مجموع الأرقام التي بين قوسين = ١٦

ويكون متوسط الانحرافات = ١٦ ÷ ٤ = ٤

من الواضح أنه تم تجاهل الإشارات السالبة، وتم جمع الفروق بين القيم والوسط الحسابي ، ثم القسمة على ٤

أما في حالة وجود البيانات في جدول تكراري، فإن حساب متوسط الانحرافات يقوم على منتصفات الفئات والوسط الحسابي.

مثال: الجدول الآتي يوضح توزيع عينة من القائمين بالاتصال في الصحف حسب مدة الخبرة في العمل بالسنوات:

مدة الخبرة	ك
١- ٣	٢٠
٤-٨	٣٤
٩- ١٣	١٦
المجموع	**٧٠**

لحساب متوسط الانحرافات ، نبدأ بتحديد مراكز الفئات (ص)، وضربها في التكرارات (ك)؛ ومن ثمّ الحصول على الوسط الحسابي (م)، ثم حساب الفروق بين الوسط الحسابي ومراكز الفئات، وضرب هذا الفارق في التكرارات الأصلية، فنحصل على الجدول الآتي:

فئات المدة الزمنية	ك	ص	ك × ص	ص - م	ك (ص - م)
١- ٣	٢٠	٢	٤٠	٤	٨٠
٤-٨	٣٤	٦	٢٠٤	صفر	صفر
٩- ١٣	١٦	١١	١٧٦	٥	٨٠
المجموع	**٧٠**		**٤٢٠**		**١٦٠**

إن الوسط الحسابي = ٤٢٠÷ ٧٠ = ٦ وهذه القيمة (٦) تم طرحها من قيم (ص)، كما هو موضح بالعمود الخامس وهو (ص- م)، مع ملاحظة أنه تم تجاهل الإشارات السالبة.

ويكون متوسط الانحرافات أو الانحراف عن الوسط = ١٦٠ ÷ ٧٠ = ٢.٣

ولتوضيح هذا الجدول فإن:

- مركز الفئة الأولى مثلًا = (١+ ٣÷٢) = ٢ (وهذا موضح في العمود الثالث) كما أن عدد تكرارات هذه الفئة هو (٢٠)، وبالتالي فإن ك × ص لهذه الفئة يساوي (٢٠×٢= ٤٠)، وهذا موضح في العمود الرابع، وهكذا في بقية الفئات ، ويكون مجموع ك × ص لجميع الفئات هو مج (ك ص) وهو يساوي (٤٢٠)

- بقسمة هذا المجموع على مجموع التكرارات نحصل على الوسط الحسابي؛ أي ٤٢٠ ÷ ٧٠ = ٦

- الحصول على الفروق وذلك بطرح قيم (ص) من الوسط الحسابي ، فإذا كانت قيمة (ص) للفئة الأولى مثلًا هي (٤)، والوسط الحسابي هو (٦)، فإن (ص- م) لهذه الفئة هو (٤- ٦)= ٢ وهكذا لكل الفئات كما هو واضح في العمود الخامس الذي عنوانه (ص- م)، وقد تم تجاهل الإشارات السالبة.

- ضرب هذه الفروق في التكرارات الأصلية لكل فئة، كمثال على ذلك، فإن التكرارات الأصلية للفئة الأولى هي (٢٠) والفرق الخاص بها هو (٤)، وبضرب التكرارات الأصلية في الفرق الخاص بهذه الفئة، فإن (٢٠ ×٤= ٨٠) وهكذا لكل الفئات، كما هو واضح في العمود السادس أي الأخير بالجدول.

- جمع ناتج عملية الضرب هذه لكل الفئات (ومن الجدول نتبين أن مجموع ناتج ضرب الفروق في التكرارات الأصلية = ١٦٠)

- متوسط الانحرافات = ١٦٠ ÷ ٧٠ = ٢.٣

غير أن متوسط الانحرافات يتجاهل الإشارات السالبة بدون منطق علمي، ولذلك فإنه غير شائع الاستخدام في البحوث العلمية كما سبقت الإشارة.

خامسًا:معامل الاختلاف

يستخدم معامل الاختلاف (Variation Coefficient) للمقارنة بين توزيعين من حيث التشتت، كأن يريد الباحث مقارنة توزيع عينة من المبحوثين حسب معدل قراءة الصحف اليومية، وتوزيع هؤلاء المبحوثين حسب الدرجة على مقياس المعرفة بقضايا البيئة، ويشيع استخدام معامل الاختلاف في المقارنة بين ظاهرتين مختلفتين في وحدة القياس، كالمقارنة بين اهتمام صحيفة يومية معنية بقضية تنظيم الأسرة واهتمام قناة تليفزيونية معنية بالقضية نفسها. هنا نلاحظ أن وحدة قياس الاهتمام في الصحيفة هي السنتيمتر المربع، أما وحدة قياس الاهتمام في القناة التليفزيونية، فهي الدقيقة؛ أي أن وحدة القياس مختلفة في الحالتين. ويمكن حساب معامل الاختلاف بأكثر من طريقة أهمها: استخدام الوسط الحسابي والانحراف المعياري، وكذلك باستخدام الانحراف الرُّبيعي والوسيط، وأيضًا باستخدام الرُّبيع الأول والرُّبيع الثالث:

١- حساب معامل الاختلاف باستخدام الوسط الحسابي والانحراف المعياري:

يمكن حساب معامل الاختلاف باستخدام الوسط الحسابي والانحراف المعياري، وذلك بموجب القانون:

$$\text{خ} = \frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}} \times ١٠٠$$

فإذا كان متوسط درجة المبحوثين على مقياس استخدام مواقع التواصل الاجتماعي هو ٣٥ درجة، بانحراف معياري ٣.٤، فإن معامل الاختلاف =

م خ = (٣.٤ ÷ ٣٥) × ١٠٠ = ٩.٧%

وإذا كان متوسط درجة المبحوثين على مقياس استخدام التليفزيون هو (٤٦.٥) درجة بانحراف معياري ٢.٨، فإن معامل الاختلاف = (٢.٨ ÷ ٤٦.٥) × ١٠٠ = ٦.٠٢

وهذا يعني أن استخدام التليفزيون أقل تشتتًا مقارنة بمواقع التواصل الاجتماعي.

٢- حساب معامل الاختلاف باستخدام الانحراف الرُّبيعي والوسيط:

يمكن حساب معامل الاختلاف باستخدام الوسيط والانحراف الربيعي، وذلك بموجب القانون:

$$\text{خ} = \frac{\text{الانحراف الرُّبيعي}}{\text{الوسط}} \times ١٠٠$$

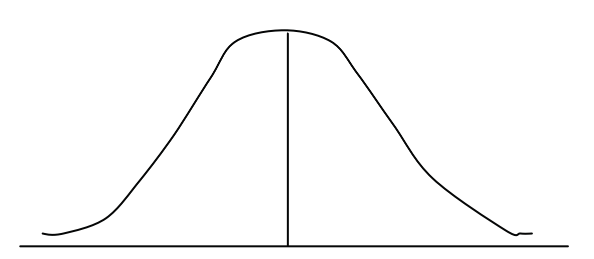
٣- حساب معامل الاختلاف باستخدام الرُّبيع الأول والرُّبيع الثالث:

يتم حساب معامل الاختلاف باستخدام الرُّبيع الأعلى والرُّبيع الأدنى، وذلك بموجب القانون:

$$\text{م خ} = \frac{\text{الرُّبيع الثالث} - \text{الرُّبيع الأول}}{\text{الرُّبيع الثالث} + \text{الرُّبيع الأول}} \times ١٠٠$$

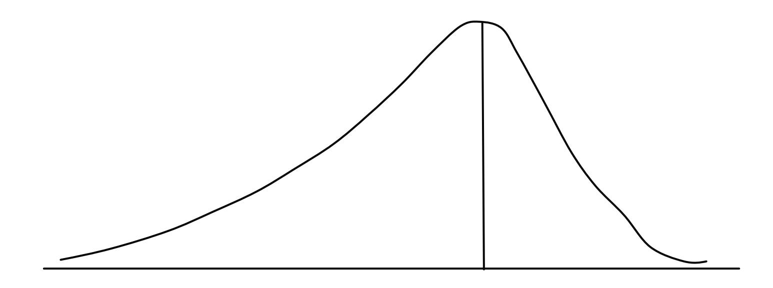
سادسًا:معامل الالتواء

الالتواء (Skew ness) هو ميل منحنى التوزيع ناحية اليمين أو ناحية اليسار، وهذا عكس اعتدالية المنحنى (عدم ميله إلى ناحية معينة)، والتوزيع المعتدل يعني أن الوسط الحسابي= الوسيط = المنوال، وهذا يعكسه المنحنى التالي:



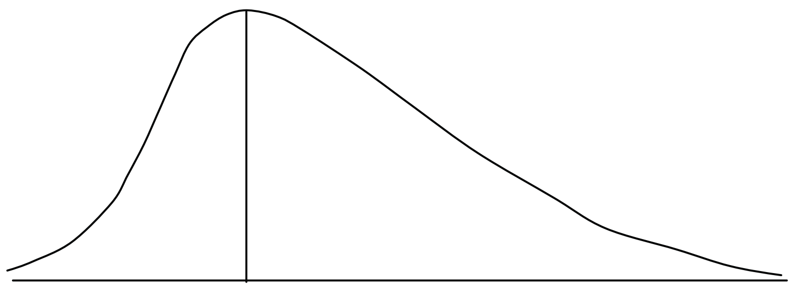
فهذا توزيع اعتدالي (مثالي)، ومن الواضح أن العمود الساقط من القمة على المحور الأفقي يقسم هذا المحور(قاعدة المنحنى) إلى قسمين متساويين.

أما إذا كان المنحنى مائلًا ناحية اليمين، فإنه يكون على الشكل التالي:



فهذا المنحنى يعني أن التوزيع موجب الالتواء، ومن الواضح أن العمود الساقط من القمة على المحور الأفقي يقسم قاعدة المنحنى إلى قسمين غير متساويين؛ حيث يميل باتجاه اليمين، وهذا يدل على ميل التوزيع إلى القيم الأعلى.

على النقيض من ذلك، فإن منحنى التوزيع قد يكون مائلًا ناحية اليسار، بمعنى أنه يتخذ الشكل التالي:



فهذا المنحنى يعني أن التوزيع سالب الالتواء، ويلاحظ أن العمود الساقط من القمة على المحور الأفقي يقسم قاعدة المنحنى إلى قسمين غير متساويين؛ حيث يوجد العمود في الجزء الأيسر، وهذا يدل على ميل التوزيع إلى القيم الأقل.

وتتراوح قيمة معامل الالتواء ما بين (-٣) إلى (+٣)، وفي التوزيع الاعتدالي، فإن قيمة الالتواء تساوي (صفر)، وكلما اقتربت قيمة عامل الالتواء من الصفر، دلّ ذلك على مثالية (اعتدالية) التوزيع، ويكون الالتواء سالبًا بشدة، كلما اقتربت قيمته من (-٣)، بينما يكون موجبًا بشدة، كلما اقتربت قيمته من (+٣).

وتتعدد المعادلات التي بواسطتها نعرف قيمة الالتواء، وإن كانت تعتمد على مقاييس النزعة المركزية (الوسط، الوسيط، المنوال) وكذلك على الانحراف المعياري، ومن أبرز المعادلات المستخدمة في حساب الالتواء تلك المعادلة:

$$\frac{\text{٣ (الوسط الحسابي – الوسيط)}}{\text{الانحراف المعياري}}$$

فإذا كان الوسط الحسابي هو (٢٨) والوسيط هو (٣٠) والانحراف المعياري (٤)، فإن الالتواء يساوي:

$$\frac{(30-28)3}{4} = \text{-١.٥}$$

أي أن التوزيع سالب الالتواء بدرجة متوسطة.

ويمكن استخدام المنوال بدلًا من الوسيط في المعادلة السابقة، ويكون الالتواء يساوي:

$$\frac{\text{٣ (الوسط الحسابي – المنوال)}}{\text{الانحراف المعياري}}$$

بوجه عام، فإن التوزيع يكون سالب الالتواء إذا كان الوسط الحسابي يقل بفروق واضحة عن الوسيط، أو المنوال، بينما يكون الالتواء موجبًا، إذا كان الوسط الحسابي يزيد بفروق ملموسة عن الوسيط أو المنوال.

سابعًا:الدرجات المعيارية

الدرجات المعيارية(Standardized Score) هي تحويل الدرجات الخام (Row Score) إلى انحرافات معيارية كوحدة قياس، فإذا افترضنا أن المبحوث قد حصل على درجة مقدارها (٢٤) مثلًا على مقياس استخدام التليفزيون، فإن هذه الدرجة هي الدرجة الخام، ويتم تحويل هذه الدرجة الخام إلى درجة معيارية - والمعروفة اصطلاحًا- (Z Score) أو الدرجة الزائية بموجب المعادلة:

$$Z = \frac{\text{س} - \overline{\text{س}}}{\text{ع س}}$$

حيث Z تعني الدرجة المعيارية، أما (س) فتعني درجة الفرد، بينما ($\overline{\text{س}}$) تعني متوسط درجة المجموعة أو العينة التي ينتمي إليها الفرد، أما (ع س) فتعني الانحراف المعياري لمتوسط درجة المجموعة أو العينة، فإذ افترضنا أن أحد المبحوثين حصل على ١٨ درجة على مقياس الاستخدامات والإشباعات، وكان متوسط درجة المجموعة ككل على هذا المقياس ١٢ بانحراف معياري ٣، فإن الدرجة المعيارية لهذا المبحوث تكون:

$$= \frac{18 - 12}{3} \text{ ٢}$$

وإذا كان أحد المبحوثين حصل على ١٢ درجة على مقياس الاستخدامات والإشباعات، وكان متوسط درجة العينة أو المجموعة على هذا المقياس ١٨ بانحراف معياري ٣، فإن الدرجة المعيارية لهذا المبحوث تكون:

$$\frac{12-18}{3} = -2$$

أي أن قيمة الدرجات المعيارية قد تكون موجبة كما قد تكون سالبة، فإذا كانت الدرجة المعيارية لأحد المبحوثين تساوي(٢)، فإن ذلك يعني أن درجته تزيد عن متوسط درجة العينة بمقدار ٢ وحدة انحراف معياري، أما إذا كانت درجة أحد المبحوثين تساوي(-٢)، فإن ذلك يعني أن درجته تقل عن متوسط درجة العينة بمقدار ٢ وحدة انحراف معياري.. وهكذا.

وباستخدام الدرجات الزائية يمكننا مقارنة درجات على مقياسين مختلفين،كما يمكننا مقارنة متوسط درجة على مقياس معين بمتوسط درجة على مقياس آخر مختلف مهما كانت قيمة الوسط الحسابي أو قيمة الانحراف المعياري؛ لأن الدرجات المعيارية توحد متوسطات درجات المقاييس؛ إذ إن متوسط الدرجات المعيارية في كل الأحوال يساوي الصفر، كما أن الانحراف المعياري لها يساوي الواحد الصحيح.

أما فيما يخص الدرجة التائية (T. Score)، أو (T. Norms)، فإنها سميت كذلك نسبة إلى كل من ثورنديك وتيرمان (كلاهما اسمه يبدأ بحرف T)، والدرجات التائية هي درجات معيارية لتوزيع معين متوسطه الحسابي خمسون وانحرافه المعياري عشرة (م = ٥٠، ع= ١٠) كما أن الدرجات التائية هي في حقيقتها توزيع جديد للدرجات الزائية، هذا التوزيع الجديد يكون أوسع نطاقًا مما يتيحه التوزيع الطبيعي، وبتوضيح أكثر، فإن المساحة التي تقع تحت المنحنى الاعتدالي المعياري تنحصر بين (-٣) و (+٣) وحدات انحراف معياري،وهذا مدى ضيق يفرض استخدام الكسور عند مقارنة الدرجات، وللتغلب على هذه المشكلة يتم تحويل الدرجات المعيارية الزائية إلى درجات معيارية تائية، متوسطها (٥٠) وانحرافها المعياري (١٠)؛ أي أننا نضرب الدرجة الزائية في ١٠ ثم نضيف ٥٠ إلى الناتج، وهكذا نتخلص من العلامات السالبة ومن الأرقام الصغيرة التي

نقارنها، ويصبح متوسط الدرجات التائية مساويًا ٥٠، أما انحرافها المعياري فيساوي ١٠؛ أي أننا بهذا المعنى عدلنا الوسط الحسابي (فبعد أن كان يساوي الصفر، أصبح يساوي ٥٠)، كما عدلنا الانحراف المعياري (فبعد أن كان يساوي الواحد الصحيح أصبح يساوي ١٠)، وعلى وجه التحديد، فإن الدرجة التائية تكون بموجب المعادلة:

T = (١٠ × الدرجة الزائية) + ٥٠ أي أن :

T. Score = 10 z Score + 50

ومعنى ذلك أن حساب الدرجة التائية (T. Score) يتطلب معرفة الدرجة الزائية أولًا، فإذا كانت الدرجة الزائية (Z. Score) لأحد المبحوثين هي ٢، فإن الدرجة التائية له تكون :

(١٠×٢) + ٥٠ = ٧٠

وإذا كانت الدرجة الزائية هي – ٥، فإن الدرجة التائية تكون:

(١٠×-٥) + ٥٠ = صفر

وعندما تكون الدرجة الزائية تساوي الصفر،فإن الدرجة التائية تكون:

(١٠× صفر) + ٥٠ = ٥٠

وهكذا يمكننا حساب الدرجات المعيارية التائية (T. Score) المقابلة للدرجات المعيارية الزائية.

خلاصة الفصل الثاني

تناول هذا الفصل عملية إعداد البيانات والمعالجة الإحصائية الوصفية، وتتلخص موضوعاته الأساسية في:

- يقصد بتجهيز البيانات إعدادها بالطريقة المناسبة؛ بحيث يمكن معالجتها إحصائيًا معالجة سليمة تقابل أهداف البحث، وتتمثل الإجراءات الأساسية لتجهيز البيانات في: المراجعة، الإدخال، التنظيم، ومن ثم وصف البيانات. ذلك أنه قبل إدخال البيانات التي ستتم معالجتها إحصائيًا تتم مراجعتها بدقة؛ ومن ثمّ إدخال هذه البيانات في البرنامج الإحصائي . أما تنظيم البيانات فيعني تكييف البيانات إحصائيًا (Statistically Adjusting the Data) بما يناسب أهداف البحث وفروضه وتساؤلاته، ويتم تنظيم البيانات من خلال إعادة التكويد (Recoding)، تكوين المتغيرات، تحديد المستويات، التحويل (Transform) ومن أهم أشكال التحويل، تحويل الدرجات الخام (Raw Scores) إلى درجات معيارية (Standardized Scores) ، التحويلات الخطية (Linear Transformations).

- المعالجة الوصفية تتضمن معرفة التكرارات والنسب (Frequencies and Percentages) وخصائص النزعة المركزية والتشتت؛ ومن ثمّ تبسيط البيانات وتلخيصها في صورة موجزة توضح أهم خواصها؛ بحيث يصل الباحث إلى وصف موضوعي للمتغيرات وتنظيمها في جداول ورسوم بيانية، ومنحنيات وأشكال توضيحية تبين معالم المتغيرات قيد البحث. ومن خلال الوصف الإحصائي يتم التمهيد للتحليل الإحصائي المناسب بما يقابل أهداف البحث.

- إن التوزيعات التكرارية تعني التكرارات والنسب المئوية (Frequencies and Percentages) للمتغيرات أيًا كانت، ويتم التعبير عنها في جداول منظمة،

والجداول التكرارية قد تكون بسيطة، كما قد تكون مزدوجة (Cross-Tabs)، وقد تكون متعددة المتغيرات (.Multivariate Variables Tabs) ، فالجدول البسيط هو الجدول الذي يتضمن تكرارات متغير واحد، أما الجدول المزدوج، فهو الجدول الذي يتضمن التوزيع حسب متغيرين معًا، أما الجداول متعددة المتغيرات (.Multivariate Variables Tabs)، فهي الجداول التي تتضمن توزيع التكرارات حسب أكثر من متغيرين (ثلاثة متغيرات أو أكثر).

- إن البيانات الكميّة لا يكون لها معنى دون تنظيمها في جداول تكرارية واضحة تبين مغزاها، ويتم تصميم الجدول التكراري بتحديد المدى (Range) الذي تشغله تلك القيم، وتحديد طول الفئة المناسب؛ ومن ثمّ تفريغ البيانات في الفئات، وقد تكون الفئات منتظمة (متساوية الطول)، كما قد تكون غير منتظمة (غير المتساوية في الطول)، كما أن الفئات قد تكون مفتوحة أو مغلقة، فالفئات المفتوحة هي الفئات التي لها بداية ونهاية محددة،أما الفئات المفتوحة فهي الفئات التي لم يتم تحديد بدايتها أو نهايتها، والجدول التكراري الأمثل هو الذي تكون فئاته مستقلة، غير متداخلة، بمعنى أن تبدأ الفئة اللاحقة من حيث انتهت السابقة عليها، وأن يكون عدد الفئات مناسبًا، مع تحاشي استخدام الفئات المفتوحة، وكذلك تحاشي الفئات غير المنتظمة قدر الإمكان.

- التوزيع التكراري المتجمع هو التوزيع الذي يوضح عدد التكرارات التي تقل أو التي تزيد على قيمة معينة. وهناك التوزيع التكراري المتجمع الصاعد، كما أن هناك التوزيع التكراري المتجمع الهابط. فالتوزيع التكراري المتجمع الصاعد هو عدد التكرارات التي تقل عن الحدالأعلى للفئات (أي أنه يبدأ بالصفر ثم بتكرارات الفئة الأولى، ثم نضيف عليها تكرارات الفئات اللاحقة)، أما التوزيع التكراري المتجمع الهابط، فهو عدد التكرارات التي تزيد على الحد

الأدنى للفئات (أي أنه يبدأ بإجمالي التكرارات، ثم ننقص منه تكرارات كل فئة من الفئات اللاحقة(من الأولى حتى الأخيرة).

- تتعدد الرسوم البيانية التي يمكن استخدامها في التعبير عن المعطيات الإحصائية، ومن أبرزها: الأعمدة، المدرج التكراري، المضلع التكراري، المنحنيات، الدوائر... إلخ، وكل شكل من هذه الأشكال يتخذ أنماطًا متعددة، ويمكن تطويعه بأساليب متنوعة لتمثيل المعطيات الإحصائية.

- النزعة المركزية (Central Tendency): هي نزوع التكرارات أو القيم الكمية الخاصة بالموضوع إلى الاقتراب من أو الابتعاد عن القيمة الوسطى أو قيمة المركز، وهي القيمة التي تتجمع حولها أكثرية التكرارات. ومن أهم مقاييس النزعة المركزية: الوسط الحسابي (Mean)، الوسيط (Median)، المنوال (Mode). إن الوسط الحسابي في جوهره هو مجموع القيم مقسومًا على عددها، أما الوسيط فهو القيمة التي تتوسط القيم بحيث يكون عدد القيم الموجودة على يسار تلك القيمة مساويًا لعدد القيم الموجودة على يمينها، في حين أن المنوال هو القيمة الأكثر شيوعًا، أو الأكثر تكرارًا.

- إن مقاييس النزعة المركزية توضح تجمع البيانات، والقيم المتوسطة لها، غير أن ذلك لا يعتبر كافيًا لوصف البيانات بدقة، وبالتالي يتعين استخدام مقاييس إضافية للوصول إلى وصف دقيق للبيانات، هذا الوصف تتيحه مقاييس التشتت، ومن أمثلتها المدى الكلي، التباين والانحراف المعياري، الانحراف الرُّبيعي، متوسط الانحرافات، معامل الالتواء، معامل الاختلاف، الدرجات المعيارية.

- المدى الكلي (Total Range): هو الفرق بين أصغر قيمة وأكبر قيمة مع إضافة واحد صحيح لهذا الفرق.

- التباين (Variance): هـو متوسـط مجمـل مربعـات الانحـراف عـن الوسـط الحـسابي، أمـا الانحراف المعياري (Standard Deviation): فهـو الجـذر للتبـاين؛ أي أن التبـاين هـو مربـع الانحراف المعياري (وبالتالي فإن الانحراف المعياري هو الجـذر التربيعـي للتبـاين). ويعتـبر التباين والانحراف المعياري من أهم مقاييس التـشتت، ويعتمـدان عـلى الوسـط الحـسابي بشكل أساسي.

- يتم حساب الانحراف الرُّبيعـي (نـصف المـدى الرُّبيعـي) مـن خـلال الرُّبيـع الأول والرُّبيـع الثالث. فالرُّبيع الأول هو تلك القيمة التي تقل عنها ٢٥% من القـيم. أمـا الرُّبيـع الثالـث فهو تلك القيمة التي تقل عنها ٧٥% من القيم، علمًا بأن الرُّبيع الثاني هو الوسيط.

- متوسط الانحرافات: هو متوسط انحرافات القيم عن وسطها الحسابي، مع تجاهل الإشارات السالبة (علامة الناقص)، وهذه نقطة ضعف أساسية في هذا الأسلوب كمقياس للتشتت.

- يـستخدم معامـل الاخـتلاف (Variation Coefficient) للمقارنـة بـين تـوزيعين مـن حيـث التشتت، كأن يريد الباحث مقارنة توزيع عينة من المبحوثين حسب معدل قراءة الـصحف اليومية، وتوزيع هؤلاء المبحوثين حسب الدرجة على مقياس المعرفة بقضايا البيئة، ويشيع استخدام معامل الاختلاف في المقارنة بين ظاهرتين مختلفتين في وحدة القياس.

- الالتواء (Skew Ness): هو ميل منحنى التوزيع ناحية اليمين أو ناحية اليـسار، وهـذا عكـس اعتداليـة المنحنـى (عـدم ميلـه إلى ناحيـة معينـة)، والتوزيـع المعتـدل يعنـي أن الوسـط الحـسابي= الوسـيط = المنـوال، وتـتراوح قيمـة معامل الالتواء ما بين (-٣) إلى (+٣). وفي التوزيع الاعتدالي، فإن قيمة الالتواء

تساوي (صفر) وكلما اقتربت قيمة عامل الالتواء من الصفر، دلّ ذلك على مثالية (اعتدالية) التوزيع، ويكون الالتواء سالبًا بشدة كلما اقتربت قيمته من (-٣)، بينما يكون موجبًا بشدة كلما اقتربت قيمته من (+٣)، وتتعدد المعادلات التي بواسطتها نعرف قيمة الالتواء، وإن كانت تعتمد على مقاييس النزعة المركزية (الوسط، الوسيط، المنوال) وكذلك على الانحراف المعياري.

- الدرجات المعيارية (Standardized score): هي تحويل الدرجات الخام (Row Score) إلى درجات معيارية كوحدة قياس، وهذه الدرجات المعيارية قد تكون درجات زائية (Z. Scores) أو درجات تائية (T. Scores).

مصادر الفصل الثاني ومراجعه

(أ) مصادر ومراجع عربية:

- رجاء محمود أبو علام (١٩٩٩)، مناهج البحث في العلوم النفسية والتربوية، (القاهرة: دار النشر للجامعات).
- سعدي شاكر حمودي (٢٠٠٠)، علم الإحصاء وتطبيقاته في المجالين التربوي والاجتماعي، (عمان: مكتبة دار الثقافة للنشر والتوزيع).
- صلاح الدين محمود علام (٢٠٠٠)، تحليل بيانات البحوث النفسية والاجتماعية والتربوية، (القاهرة: دار الفكر العربي).
- عبد الحميد محمد نجم & محمد عبد الهادي المحميد (١٩٩٠)، الإحصاء الوصفي والتحليلي مع استخدام البرامج الجاهزة، (الكويت: مكتبة جامعة الكويت).
- فؤاد البهي السيد (١٩٧٨)، علم النفس الإحصائي، (القاهرة: دار الفكر العربي).

(ب) مصادر ومراجع أجنبية:

- Coolican, Hugh (2004) Research Methods and Statistics in Psychology, 4th ed. England. Abingdon. Book Point Ltd

- Hasan, Mohammed A.; Thaut, Michael H. (2004) Statistical Analysis for Finger Tapping With a Periodic External Stimulus. Perceptual and Motor Skills, Vol. 99 (2), pp. 643-66.

- Hawkins, Richard John (2005) Language Loss in Guatemala: A Statistical Analysis of the 1994 Population Census. Journal of Sociolinguistics, Vol. 9 (1), pp. 53-73.

- Hecht, Deborah (2003) Issues of Research Design and Statistical Analysis; In: Studying Service-learning: Innovations in Education

Research methodology. Billig, Shelley H.; Waterman, Alan S.; Mahwah, NJ, US: Lawrence Erlbaum Associates, Publishers, pp. 107-123.

- Ho, Hsiu-Zu; O'Farrell, Stacy L.; Hong, Sehee (2006) Developmental Research: Theory, Method, Design and Statistical Analysis.; In: Handbook of Complementary Methods in Education Research. Green, Judith L.; Camilli, Gregory; Elmore, Patricia B.; Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, pp. 207-225.

- Huckin, Thomas (2004) Content Analysis: What Texts Talk About. In: What Writing does and How it Does it: An Introduction to Analyzing Texts and Textual Practices. Bazerman, Charles; Prior, Paul; Mahwah, NJ, US: Lawrence Erlbaum Associates, Publishers, pp. 13-32.

- Huisman, Mark; Snijders, Tom A. B. (2003) Statistical Analysis of Longitudinal Network Data with Changing Composition. Sociological Methods & Research, Vol. 32 (2), pp. 28-253.

- Leedy, Paul D.&Jeanne, Ellis Ormrod (2005) Practical Research: Planning and Design. Pearson Merrill Prentice Hall.

- Li, Heng; Psychometrika (2004) A Widely Applicable Extension of the Random Effects Two-Way Layout: Its Definition and Statistical Analysis Based on Group Invariance. Vol. 69 (1), pp. 55-64.

- Meyer, Merle E. (2005) A Statistical Analysis Of Behavior. Oxford, England: Wadsworth.

- Ohtani, Hideo; Kobayashi, Masayuki (2005) Statistical Analysis of Dangerous Goods Accidents in Japan ; Safety Science, Vol. 43(5-6), pp. 287-297.

- Olatunji, Bunmi O.; Feldner, Matthew T.; Witte, Tricia H. (2004) Graduate Training of the Scientist-Practitioner: Issues in Translational Research and Statistical Analysis. ; Behavior Therapist, Vol. 27(3), pp. 5-45.

- Pearl, Judea (1971) Application of Walsh Transform to Statistical Analysis. EEE Transactions on Systems, Man, & Cybernetics, SMC-1(2), Apr 1971. pp. 111-119.

- SPSS Manuals. Different Versions. London. Sage.

- Vaus, David de (2002) Analyzing Social Science Data. London. SAGE Publications.

- Weiss, Robin S.; Remington, Roger; Ellis, Stephen R. (1989) Sampling Distributions of the Entropy in Visual Scanning. Behavior Research Methods, Instruments & Computers, Vol. 21(3), pp. 348-352.

- Zimmerman, Donald W.; Zumbo, Bruno D. (2006) Can Percentiles Replace Raw Scores in the Statistical Analysis of Test Data?; Educational and Psychological Measurement, Vol. 65(4), pp. 616-638

الفصل الثالث

مقاييس العلاقة والارتباط

تمهـــيد

تقوم معاملات الارتباط(Correlation Coefficients) بـدور أسـاسي في مختلـف البحـوث العلميـة، ولا يكاد يخلو بحث إمبيريقـي مـن هـذه المعـاملات، سـواء بشـكل مبـاشر أو غـير مبـاشر، بـل إن معظـم الأساليب الإحصائية تقوم على معاملات الارتباط.وقد لوحظ أن هناك بعض المثالب في اختيـار وتطبيـق معاملات الارتباط، خاصة من حيث اختيـار الـنمط الـذي يناسـب المتغـيرات المطلـوب تـقصي الارتبـاط بينها.. الفصل الحالي يلقي الضوء على الجوانب الأساسية لمقاييس العلاقة والارتباط مـن خـلال توضـيح معنـى الارتبـاط وخصائصـه، والاختيـار الصـحيح لمعـاملات الارتبـاط،وأهم معـاملات الارتبـاط شـائعة الاستخدام في البحوث الاجتماعية مع توضيح معامل الارتباط المتعدد، وكذلك الارتباط الجـزئي باعتبـاره أحد أدوات الضبط الإحصائي الهامة، سواء في الدراسات المسحية أو التجريبية.

أولًا: معنى الارتباط وخصائصه

الارتباط (Correlation) بمعناه الإحصائي: هو اقتران التغير في ظاهرة معينة بالتغير في ظاهرة أخـرى، بمعنى أنه إذا تغيرت الظاهرة (س) تتغير الظاهرة (ص)، وهذا التغـير يعـبر عنـه بمعامـل الارتبـاط، وفي المتغيرات الكمية، فإن الارتباط من حيث طبيعته قد يكون موجبًا أو سالبًا:

- فالارتباط الموجب يعني زيادة في الظاهرة (س) عندما تزداد الظاهرة (ص)، أو زيادة في الظاهرة (ص) عندما تزداد الظاهرة (س).

- أما الارتباط السالب فيعني نقصان في الظاهرة (س) عندما تـزداد الظـاهرة (ص)، أو نقصـان في الظاهرة (ص) عندما تزداد الظاهرة (س).

أما من حيث الاتجاه (Direction)، فإن الارتباط قد يكون طرديّا أوعكسيّا، فالارتباط الطـردي يعنـي تغير الظاهرتين في اتجاه واحد (بالزيادة أو النقصان)، أما الارتباط العكسـي فيعنـي تغـير الظـاهرتين في اتجاهين متناقضين كزيادة الظاهرة (س) عندما تتناقص الظاهرة (ص)، أو تناقص الظـاهرة (س) عنـدما تزيد الظاهرة (ص).

وتتراوح قيمة معامل الارتباط بين + 1 إلى – 1؛ أي من الارتباط الموجب التام إلى الارتباط السـالب التام، ويصل الارتباط إلى نهايته العظمى عندما يقترن تغير المتغـير الأول اقترانًـا تامّـا بـالتغير في المتغـير الثاني، هذا الارتباط التام قد يكون موجبًا أو سالبًا. فقـد يتبـين مـن الدراسـة مـثلًا اقـتران زيـادة وقـت مشاهدة التليفزيون بزيادة وقت الفراغ؛ بحيث يظل ترتيب الأفراد بالنسبة للمتغيرين ثابتًـا، فكلـما زاد وقت مشاهدة التليفزيون زاد وقت الفراغ، وكلما زاد وقت الفراغ زاد وقت مشاهدة التليفزيون؛ ونظرًا لأن مشاهدة التليفزيون تزيد بنفس وحدات زيادة وقت الفراغ، فإن ترتيب الأفراد يظل كما هو- ثابتًا- حسب كلا المتغيرين، وهنا يكون هناك ارتباط موجب تام؛ أي أن قيمته تكون مساوية للواحد الصـحيح (+1)، أما الارتباط التام السالب، فقد يتضح من بحوث التليفزيون مثلًا، اقتران زيادة الوقـت المنقضــي في مشاهدة التليفزيون بنقصان الوقت المنقضــي في التحـدث مـع أفـراد الأسرة؛ بحيـث يكـون ترتيـب الأفراد واحدًا في المتغيرين (الوقت المنقضي في مشاهدة التليفزيون والوقـت المنقضــي في التحـدث مـع أفراد الأسرة)، وتكون قيمة الارتباط التام السالب مساوية لناقص الواحد الصحيح (-1).

ولا تتأثر قيمة معامل الارتباط بزيادة أو نقصان تكرارات المتغيرين بكميات ثابتة، فقد يتضح من إحدى الدراسات مثلًا أن الارتباط بين الوقت اليومي المنقضي في مشاهدة التليفزيون والوقت اليومي المنقضي في استخدام (Facebook) هو 0.28، فإذا أضفنا ساعة من الوقت اليومي المنقضي في مشاهدة التليفزيون وأضفنا ساعة من الوقت اليومي المنقضي في استخدام (Facebook) وحسبنا معامل الارتباط مرة أخرى، فإننا نجده 0.28 أي أنه ظل ثابتًا، فهذه الإضافة لا تؤثر في ترتيب الأفراد حسب المتغيرين ويبقى التغير الاقتراني القائم بين المتغيرين كما هو. المنطق نفسه إذا طرحنا - أي أنقصنا- عددًا ثابتًا من قيم المتغيرين، فهذا النقصان لا يؤثر في الترتيب، ويمكن أن نستعين بهذه الفكرة في تبسيط العمليات الحسابية، وذلك بإضافة أو بطرح عدد ثابت من قيم المتغيرات التي تحسب معاملات ارتباطها.

ثانيًا: اختيار معامل الارتباط المناسب

تتعدد معاملات الارتباط، ويتوقف اختيار معامل الارتباط المناسب على نوعية المتغيرات المطلوب حساب الارتباط بينها؛ أي ما إذا كانت تلك المتغيرات اسمية (Nominal) أو رتبية (Ordinal) أو فاصلة (Interval) وهي المتغيرات التي يشيع استخدامها في العلوم الإنسانية. أي أن خصائص المتغيرات هي التي تحدد معامل الارتباط المناسب، وبجانب طبيعة المتغيرات - أي ما إذا كانت اسمية أو رتبية أو فئوية (فترية) - هناك عوامل أخرى تحدد معامل الارتباط المناسب، منها عدد التقسيمات التي يضمها كل متغير، فهناك متغيرات ذات تقسيم ثنائي، منها مثلًا: متغير الجنس (ذكور & إناث)، أو متغير مشاهدة التليفزيون (يشاهد & لا يشاهد)، أو متغير استخدام الإنترنت (يستخدم & لا يستخدم)... إلخ، وهناك متغيرات ذات تقسيم ثلاثي، منها مثلًا: متغير قراءة الصحف أو مشاهدة التليفزيون أو سماع الراديو (بانتظام، أحيانًا، نادرًا)، وقد يتضمن المتغير أكثر من ثلاثة تقسيمات حسب طبيعة الدراسة، كأن يوجد للبند الواحد أربع استجابات أو

أكثر يكون على المبحوث اختيار واحدة منها. ومن العوامل الأخرى التي تحدد معامل الارتباط المناسب موقع المتغير في المعالجة (كمتغير مستقل X أو كمتغير تابع Y)، وعلى الرغم من ذلك تبقى طبيعة المتغيرات هي العامل الحاسم في تحديد؛ أي معامل الارتباط تستخدم.

وحسب كل من برايمان وكرامر (Bryman & Cramer, 1999, p. 201)، فإن استخدام مقاييس العلاقة والارتباط المناسبة يتحدد حسب طبيعة المتغيرات على النحو الآتي:

(1) متغيران اسميان (Nominal – Nominal) يتم استخدام معامل التوافق مع كا^2 تربيع (Contingency Coefficient in Conjunction with Chi-Square) لاختبار الدلالة الإحصائية، ولمعرفة شدة الارتباط بين المتغيرين في هذه الحالة يمكن استخدام معامل كرامر (Cramer's V).

(2) متغيران رتبيان (Ordinal - Ordinal) يمكن استخدام معامل ارتباط سبيرمان، وكذلك معامل كندال (Kendall's Tau).

(3) متغيران فئويان (فتريان) (Interval - Interval) يتم استخدام معامل ارتباط بيرسون، ومعامل الانحدار لتقدير شدة وطبيعة العلاقة.

(4) متغيران ثنائيان (Dichotomous - Dichotomous) أيّا كانت طبيعتهما اسمية أو رتبية أو فئوية (فترية)، يتم استخدام التوافق مع كا^2 (Contingency Coefficient in Conjunction with Chi-Ssquare) لاختبار الدلالة الإحصائية، ولمعرفة شدة الارتباط بين المتغيرين في هذه الحالة يمكن استخدام معامل Phi.

(5) متغير فئوي (فتري) ومتغير رتبي (Interval- Ordinal) إذا كان المتغير الرتبي يضم أكثر من تقسيمين يمكن استخدام معامل ارتباط سبيرمان، وكذلك معامل كندال (Kendall's Tau)، كما يمكن استخدام معامل التوافق مع كا^2 إذا كان المتغير الرتبي والمتغير الفئوي (الفتري) كلاهما ثنائي التقسيم. وإذا كان المتغير الفئوي (الفتري) محددًا كمتغير تابع يمكن استخدام المتوسطات (المقارنة بين المجموعات الرتبية من حيث المتغير الفئوي (الفتري)) ومعرفة معنوية الفروق بين هذه المجموعات.

(6) متغير فتري مع متغير اسمي أو مع أي متغير آخر ثنائي رتبي (Interval- Nominal or Dichotomous) يمكن استخدام معامل التوافق مع كا^2 إذا كان المتغير الفتري تم تنظيمه في مجموعات مع عدم ضرورة لأن نتعامل معه كمتغير تابع (Dependent) أما إذا كان المتغير الفتري تم تحديده كمتغير تابع، فيمكن استخدام المتوسطات (المقارنة بين مجموعات المتغير الاسمي من حيث المتغير الفئوي (الفتري)) ومعرفة معنوية الفروق بين هذه المجموعات.

(7) متغير اسمي ومتغير رتبي (Nominal- Ordinal)، يتم استخدام معاملات الارتباط السابق الإشارة إليها مع المتغيرين الاسميين؛ حيث يمكن استخدام معامل التوافق مع كا^2 (Contingency Coefficient in Conjunction With Chi-Square) لاختبار الدلالة الإحصائية، واستخدام معامل كرامر (Cramer's V) لمعرفة شدة الارتباط بين المتغيرين.

وكثيرًا ما تستخدم مقاييس لابارامترية (Non-Parametric) لرصد وتحليل العلاقة والارتباط بين المتغيرات المختلفة؛ للتحقق من الفروض الارتباطية، وقد أخذ ديفيد دي فايس (David de Vaus.2002) مسألة الإحصاء اللابارامتري في الاعتبار، مؤكدًا على أن

معامل الارتباط المناسب يتحدد في ضوء طبيعة وعدد تقسيمات المتغيرين المراد قياس الارتباط بينهما، وخلص إلى رؤية واضحة تتمثل ملامحها الأساسية فيما يوضحه الجدول الآتي:

المتغيران	التقسيم		
	كلاهما ثنائي التقسيم	كلاهما تقسيم أكثر من ثنائي	أحدهما تقسيم ثنائي والآخر تقسيم أكثر من ثنائي
اسميان	Phi- Yule' Q Lambda- Goodman and Kruskall's Tau	Cramer's V- Lambda- Goodman and Kruskall's Tau	Cramer's V- Lambda- Goodman and Kruskal' Tau
رتبيان	Gamma - Somers's Kendal's Tau	Spearman' Rho- Kendal's Tau	Gamma - Somers's Kendal's Tau
فتريان	Pearson- Phi	Pearson - Eta	Pearson
اسمي & رتبي	Theta (Wilcoxon)	Theta (Wilcoxon)	Theta (Wilcoxon)
اسمي & فتري	Pearson- Phi	Eta Theta	Eta – Theta Pearson
رتبي & فتري	Pearson- Phi- Biserial	Jasper's Multi-Serial	Pearson Jasper's Multi-Serial

بوجه عام، فإن استخدام معامل الارتباط المناسب يعتمد على طبيعة المتغيرات الخاضعة للدراسة، وهذا المثال إنما هو للتوضيح وتأكيد الفكرة؛ لأن المسألة لا تتوقف عند مجرد معامل الارتباط، وإنما تشمل العديد من أساليب التحليل الإحصائي.

ثالثًا: أمثلة لمقاييس العلاقة والارتباط بين متغيرين

سبقت الإشارة إلى أن اختيار معامل الارتباط المناسب يتوقف على نوعية المتغيرات المطلوب حساب الارتباط بينها؛ أي ما إذا كانت تلك المتغيرات اسمية (Nominal) أو رتبية (Ordinal) أو فاصلة فئوية (فترية) (Interval) وهي المتغيرات التي يشيع

استخدامها في العلـوم الإنسـانية.إن الارتبـاط بـين متغـيرين (Two variables) قـد يكـون بـين: متغـيرين اسميين، أو متغيرين رتبيين، أو متغيرين فتريين، كما قد يكون الارتباط بين متغير اسمي والآخر رتبـي، أو بين متغير اسمي والآخر فتري، أو بين متغـير رتبـي والآخـر فـتري، ولكـل ذلـك معـاملات الارتبـاط التـي تناسبه، وفيما يلي أمثلة توضيحية لهذه المعاملات:

(1) ارتباط بيرسون:

يستخدم ارتباط بيرسون لحساب الارتباط بين متغيرين كلاهما من النوع الفتري (مثـال ذلـك عنـدما نريد حساب الارتباط بين الوقت المنقضي في قراءة الصحف والوقت المنقضـي في مشـاهدة التليفزيـون، أو الارتباط بين السن وعـدد سـاعات اسـتخدام الإنترنـت، أو بـين الـوزن وعـدد سـاعات الجلـوس أمـام التليفزيون أو الكومبيوتر... إلخ)، ويعتبر معامل ارتبـاط بيرسـون مقياسًـا معياريّـا للعلاقـة، بمعنـى أنـه يدخل في حسابه الوسط والانحراف المعياري لكل من مجموعتي الدرجات المراد إيجاد العلاقـة بيـنهما، وهذا يعني أن أي تحويل خطي لإحدى مجموعتي الدرجات لا يؤثر في قيمـة معامـل ارتبـاط بيرسـون، وبذلك لا يكون لوحدة القياس أهمية عند إيجاد معامل الارتبـاط. ويعـد معامـل ارتبـاط بيرسـون أكـثر أنواع معاملات الارتباط استخدامًا في البحوث بمجالاتها المختلفة، بـل إن الكثـير مـن معـاملات الارتبـاط الأخرى تعتبر حالات خاصة من معامل ارتباط بيرسون.

تتعدد صيغ الحصول على معامل ارتباط بيرسون بين متغيرين، ومن أبرز هذه الصـيغ حسـاب هـذا الارتباط بين متغيرين (س)، (ص) من الدرجات الخام مباشرة، ويكون ذلك بموجب المعادلة:

$$\frac{\text{ن مج س ص} - (\text{مج س})(\text{مج ص})}{\sqrt{\{\text{ن مج س}^2 - (\text{مج س})^2\}\{\text{ن مج ص}^2 - (\text{مج ص})^2\}}}$$

حيث:

ن: عدد مفردات العينة

مج س ص: مجموع حاصل ضرب قيم المتغير (س) في قيم المتغير (ص)

مج س: مجموع قيم المتغير (س)

مج ص: مجموع قيم المتغير (ص)

مج $س^2$: مجموع مربعات قيم المتغير (س)

مج $ص^2$: مجموع مربعات قيم المتغير (ص)

مثال توضيحي:

الجدول الآتي يوضح درجات عينة من طلاب الماجستير الإلكتروني حسب عدد الأفلام الوثائقية التي شاهدوها في قناة الجزيرة (متغير س)، وقناة المنوعات (متغير ص):

أفلام قناة الجزيرة (س)	أفلام قناة المنوعات (ص)
7	6
6	7
8	9
4	7
5	6
7	8
5	6
5	6
7	9
7	8

فإذا أردنا إيجاد معامل ارتباط بيرسون بين عـدد الأفـلام الوثائقيـة التـي شـاهدها الطـلاب في قنـاة الجزيرة (متغير س)، وعدد الأفلام التي شاهدوها في قناة المنوعات (متغير ص) فإنه يتعين الحصول على كل المعطيات التي يتطلبها تطبيق المعادلة السابقة، وهذه المعطيات هي: عدد الطلاب، مج س ص، مج س، مج (س)^2، مج ص، مج (ص)^2، مج س^2، مج ص^2. ويتم الحصول على هذه المعطيات مـن البيانـات الأصلية، وتنظيمها في جدول واضح على النحو الآتي:

أفلام قناة الجزيرة (س)	أفلام قناة المنوعات (ص)	س^2	ص^2	س × ص
7	6	49	36	42
6	7	36	49	42
8	9	64	81	72
4	7	16	49	28
5	6	25	36	30
7	8	49	64	56
5	6	25	36	30
5	6	25	36	30
7	9	49	81	63
7	8	49	64	56
مج س = 61	**مج ص = 72**	**س^2 = 387**	**ص^2= 532**	**مج س ص= 449**

من هذا الجدول يتضح لنا الآتي:

- قيمة ن تساوي 10

- مج س ص= 449
- مج س = 61 وبالتالي فإن مج $(س)^2$ = 3721
- مج ص= 72 وبالتالي فإن مج $(ص)^2$ = 5184
- مج $س^2$ = 387
- مج $ص^2$ = 532

باستخدام هذه المعطيات الإحصائية، فإن معامل ارتباط بيرسون يكون بموجب المعادلة المذكورة وهي:

$$\frac{\text{ن مج س ص} - (\text{مج س})(\text{مج ص})}{\sqrt{\{\text{ن مج س}^2 - (\text{مج س})^2\}\{\text{ن مج ص}^2 - (\text{مج ص})^2\}}}$$

وبالتعويض في هذه المعادلة يكون:

$$\frac{١٠ \times ٤٤٩ - (٦١)(٧٢)}{\sqrt{\{(١٠ \times ٣٨٧) - (٣٧٢١)\}\{(١٠ \times ٥٣٢ - ٥١٨٤)\}}}$$

$$= \frac{٤٤٩٠ - ٤٣٩٢}{\sqrt{١٤٩ \times ١٣٦}}$$

$$\frac{٩٨}{١٤٢,٣٥} = 0.688$$

أي أن معامل ارتباط بيرسون بين المتغيرين س، ص هو 0.688 ومن الواضح أنه ارتباط موجب، وهذا يعني أنه كلما زادت مشاهدة الأفلام الوثائقية عبر قناة الجزيرة، زادت مشاهدة الأفلام الوثائقية عبر قناة المنوعات.

ويمكن حساب معامل ارتباط بيرسون للبيانات غير المبوبة (سواء من خلال استخدام الدرجات المعيارية، أو باستخدام الانحرافات عن الوسط الحسابي، أو باستخدام

الدرجات الخام مباشرة، أو باستخدام الفروق بين الدرجات الخام)، كما يمكن حساب معامل ارتباط بيرسون للبيانات المبوبة، وتبدو أهمية ذلك عندما تشتمل البيانات على عدد كبير من أزواج القيم؛ حيث يمكن تبويب (جدولة) تلك القيم في جدول تكراري مزدوج (Two-Way Frequency Table)، ثم إيجاد معامل ارتباط بيرسون لهذه البيانات المبوبة باستخدام طريقة الترميز (Code Method). غير أن استخدام تلك الطريقة يتطلب تصحيح معامل الارتباط من الأخطاء الناتجة عن تبويب البيانات، ذلك أن قيمة معامل الارتباط من البيانات المبوبة تكون قيمة تقريبية؛ والسبب في ذلك يرجع إلى أننا اعتبرنا أن تكرار كل فئة يقع في مركز تلك الفئة، وكلما زاد طول الفئة، زاد بالطبع الناتج عن هذا التقريب. فإذا أراد الباحث أن يحصل على القيمة المضبوطة لمعامل ارتباط بيرسون، فعليه أن يستخدم الدرجات الخام مباشرة بدلًا من استخدام طريقة الترميز.

ويمكن تصحيح الأخطاء الناتجة عن تبويب البيانات لأي عدد من فئات كل من المتغيرين بقسمة معامل الارتباط الناتج من استخدام طريقة الترميز على مقدار ثابت يساوي عدد هذه الفئات، وقد أعد بعض علماء الإحصاء قائمة من الثوابت التي يمكن أن يستخدمها الباحث لإجراء تصحيح معامل الارتباط، عندما تبوب البيانات في فئات مختلفة السعة بالنسبة للمتغيرين (س)، (ص).

(2) معامل الاقتران:

يستخدم معامل الاقتران (Association Coefficient) في الكشف عن الارتباط بين متغيرين اسمين أو نوعيين (Nominal) كلاهما ثنائي التقسيم مثل النوع (ذكور &إناث)، قراءة الصحف (يقرأ & لا يقرأ).... وهكذا، ويتم الحصول على معامل الاقتران بواسطة المعادلة:

$$\frac{\text{أ} \times \text{د} - \text{ب} \times \text{جـ}}{\text{أ} \times \text{د} + \text{ب} \times \text{جـ}}$$

مثال: الجدول الآتي يوضح توزيع عينة قوامها (400) مفردة حسب قراءة الصحف اليومية، واستخدام مواقع التواصل الاجتماعي عبر الإنترنت:

استخدام مواقع التواصل	قراءة الصحف	
	نعم	لا
نعم	(أ) 215	(ب) 20
لا	(ج) 120	(د) 45

فإذا أردنا حساب معامل الاقتران، فإنه بموجب المعادلة المذكورة يكون:

$$0.602 = \frac{(120 \times 20) - (45 \times 215)}{(120 \times 20) + (45 \times 215)}$$

أي أن هناك ارتباطًا موجبًا قدره (0.602) بين قراءة الصحف اليومية واستخدام مواقع التواصل الاجتماعي عبر الإنترنت.

(3) **معامل ارتباط سبيرمان للرتب:**

كثيرًا ما يهدف البحث إلى قياس الارتباط بين ترتيب الأفراد أو الأشياء في صفة معينة وترتيبهم في صفة أخرى، وهنا يستخدم معامل ارتباط سبيرمان للرتب (Spearman's Coefficient of Rank Correlation)، كما يمكن استخدام هذا المعامل إذا كان أحد المتغيرين أو كلاهما من النوع الفئوي أو النسبي، وذلك بعد تحويل البيانات إلى رتب.

وفي مجال الإعلام قد يريد الباحث معرفة الارتباط بين ترتيب الأفراد من حيث الاهتمام بمشاهدة التليفزيون وترتيبهم من حيث الاهتمام بقراءة الصحف، وفي كثير من الأحيان يصعب قياس متغير ما قياسًا كميّا، ولكن يسهل تعيين رتب للصفة أو السلوك المراد دراسته عن هذا المتغير، فمثلًا إذا كان لدينا خمسة مقالات صحفية وأردنا التمييز بين

هذه المقالات من حيث إعجاب القراء بها، نجد أنه يسهل على الأفراد ترتيب تلك المقالات حسب درجة إعجابهم بها، كأن نطلب من المبحوثين إعطاء تقدير لكل مقال (ضعيف جدّا، ضعيف، متوسط، أعلى من المتوسط، ممتاز) وتكون الرتب لهذه التقديرات على التوالي هي (1، 2، 3، 4، 5)، وقد يريد الباحث معرفة ارتباط ذلك بمستوى التعليم (يقرأ ويكتب، الابتدائية، الإعدادية، الثانوية، جامعي فأعلى)، وتكون الرتب لهذه المستويات على التوالي هي (1، 2، 3، 4، 5)؛ أي أن كلا المتغيرين يأخذ رتبًا معينة، مع مراعاة أن تكون رتب المتغيرين في اتجاه واحد (من الأكبر إلى الأصغر، أو من الأصغر إلى الأكبر وذلك لكلا المتغيرين)، (فلا يصح مثلًا ترتيب أحد المتغيرين من الأصغر إلى الأكبر، وترتيب المتغير الثاني من الأكبر إلى الأصغر)، وبموجب ذلك يتم استخدام معامل ارتباط سبيرمان للرتب وفق القانون:

$$ر = 1- \frac{٦ \text{ مج ف}^{٢}}{ن (ن^{٢} - ١)}$$

حيث ر: معامل الارتباط

ن: عدد مفردات العينة، وتعبر عن أزواج البيانات (للمتغيرين)

ف: الفرق بين رتب المتغيرين (وبالتالي فإن $ف^{2}$ تعني مجموع مربعات الفروق بين رتب المتغيرين)

فإذا كان لدينا عينة قوامها 30 من الشباب، وقد حصل كل منهم على درجة معينة على مقياس الاستخدامات والإشباعات، ودرجة أخرى على مقياس المعرفة، فإن حساب معامل الرتب لسبيرمان يتم كالآتي:

- ترتيب هؤلاء المفحوصين حسب درجاتهم على المقياس الأول، وكذلك ترتيبهم حسب درجاتهم على المقياس الثاني، فيكون لكل مفحوص ترتيبان (الترتيب على المقياس الأول والترتيب على المقياس الثاني).

- حساب الفروق بين ترتيب كل مفحوص في الدرجة على المقياس الأول وترتيب المفحوص نفسه في الدرجة على المقياس الثاني، فإذا كان أحد المفحوصين قد حصل على الترتيب السابع في الدرجة على المقياس الأول وحصل على الترتيب العاشر في الدرجة على المقياس الثاني، فإن الفرق يكون 7 - 10 = -3

- تربيع الفروق بين الترتيبين وذلك لكل مبحوث في العينة، فإذا كان الفرق بين الترتيبين للمبحوث الأول هو -3، فإن تربيع هذا الفرق يكون: -3 × -3 = 9 وهكذا لكل فرد من أفراد العينة.

- جمع مربعات الفروق، فنحصل بذلك على مج ف^2.

وباستخدام المعادلة المذكورة يمكن حساب معامل ارتباط الرتب، فإذا كانت العينة تتكون من 30 شخصًا وكان مجموع مربعات الفروق هو 428 مثلًا، فإن معامل ارتباط الرتب:

$$ر = 1 - \frac{٦\,(٤٢٨)}{٣٠\,(٩٠٠-١)} = 0.905$$

مثال توضيحي:

فيما يلي درجات عينة قوامها عشرة مبحوثين على مقياس كثافة استخدام مواقع التواصل الاجتماعي (المتغير س) ودرجاتهم على مقياس المعرفة بالسياسة الداخلية (المتغير ص):

المبحوث	الأول	الثاني	الثالث	الرابع	الخامس	السادس	السابع	الثامن	التاسع	العاشر
س	15	12	13	9	10	8	14	11	16	7
ص	14	7	14	8	11	9	13	15	10	6

نلاحظ في هذا المثال أن لكل مبحوث درجة على المتغير (س)، ودرجة أخرى على المتغير (ص)، ونقطة البدء هي إعطاء رتبة لكل درجة؛ أي أن المبحوث الواحد يكون له

رتبة على المتغير (س)، ورتبة ثانية على المتغير (ص)، وذلك حسب درجته، مع مراعاة أن القيم المتكـررة تأخذ الرتبة نفسها، ثم يتم حساب الفروق بين الرتب على المتغير (س) والرتب على المتغـير (ص)، ومـن ثمّ تربيع هذه الفروق، فنحصل على الجدول التالي:

(س)	(ص)	رتب (س)	رتب (ص)	(ف)	($ف^2$)
15	14	2	2	صفر	صفر
12	7	5	8	3	9
13	14	4	2	2	4
9	8	8	7	1	1
10	11	7	4	3	9
8	9	9	6	3	9
14	13	3	3	صفر	صفر
11	15	6	1	5	25
16	10	1	5	4	16
7	6	10	9	1	1

إن إجمالي $ف^2$ وهي مربعات الفروق (العمود الأخير) يساوي (74)، أي أن (مج $ف^2$ = 74)، كما أن لدينا عشرة مبحوثين؛ أي أن قيمة ن = 10 وبالتعويض في المعادلة المذكورة نحصل عـلى معامـل ارتبـاط الرتب (سبيرمان)؛ أي أن:

$$\text{ر} = 1 - \frac{٦(٧٤)}{١٠(٩٩)} = 0.55$$

وبالرجوع إلى جدول القيم الحرجة لمعامل ارتباط الرتب لسبيرمان عند درجة حرية (10) في اختبـار ذي النهاية الواحدة نجد أن القيمة الجدولية هي (0.564)؛ أي أن القيمـة المحسـوبة أقـل مـن القيمـة الجدولية، وبالتالي لا يوجد ارتباط بين كثافة اسـتخدام مواقـع التواصـل الاجتماعـي والمعرفـة بالسياسـة الداخلية، ويمكن تفسير ذلك بأن الكثيرين مـن هـؤلاء الـذين يسـتخدمون مواقـع التواصـل الاجتماعـي بكثافة لا يستخدمونها بالضـرورة لمعرفة السياسة الداخلية للدولة، وإنمـا يسـتخدمونها لأغـراض أخـرى مثل التسلية وقضاء الوقت... إلخ.

ويشيع استخدام معامل ارتباط الرتب لسبيرمان في حالات العينات صغيرة الحجم ($n \leq 30$)، ويمكن التعرف على الدلالة الإحصائية لمعامل ارتباط الرتب باستخدام الجداول الإحصائية، وذلك عند مستوى ثقة معين بدلالة حجم العينة، فإذا كانت قيمة معامل الارتباط المحسوبة أقل من القيمة الجدولية، فإن ذلك يعني أنه لا يوجد ارتباط بين المتغيرين، أما إذا كانت قيمة معامل الارتباط المحسوبة أكبر من القيمة الجدولية، فإن ذلك يعني وجود ارتباط بين المتغيرين.

(4) معامل ارتباط جاما:

يستخدم ارتباط جاما (Gamma) أيضًا لمعرفة الارتباط بين المتغيرات الرتبية، ويشيع استخدامه بكثرة إذا كان لدينا رتبتان فقط لكل متغير، خاصة عندما تتكرر بعض القيم في المتغير الواحد، ويطلق على البيانات من هذا النوع البيانات المربوطة أو المنعقدة (Tied Data)، ويكون الترتيب لكل متغير له نطاق ضيق مما يؤدي إلى زيادة في التكرارات لنفس الرتبة ولنفس الأشخاص، وحينما يصاحب ذلك عدد كبير لأزواج قيم المتغيرين، فإنه لا يكون من المناسب استخدام معامل سبيرمان، ويستخدم معامل جاما أيضًا إذا كانت التكرارات تتوزع في خلايا الجدول بفروق كبيرة، بمعنى أن غالبية أو معظم التكرارات تتجمع في خلية واحدة أو في عدد محدود من الخلايا بينما يوجد في الخلايا الأخرى تكرارات قليلة، ويعتمد قانون معامل جاما على جدول التكرار المزدوج وعلى حالات الاتفاق والاختلاف بين أزواج القيم، ويتمثل هذا القانون في:

$$ج = \frac{ن_ق - ن_خ}{ن_ق + ن_خ}$$

حيث ج: معامل جاما

$ن_ق$: حاصل ضرب عدد حالات الاتفاق

$ن_{خ}$: حاصل ضرب عدد حالات الاختلاف

مثال:

الجدول الآتي يوضح توزيع عينة مـن المبحـوثين حسـب معـدل قـراءة الصحف ومعـدل مشـاهدة التليفزيون:

معدل مشاهدة التليفزيون	معدل قراءة الصحف	
	منخفض	متوسط
منخفض	28 (أ)	41 (ب)
متوسط	12 (ج)	23 (د)

إن معامل جاما يعتمد على تكرارات الاتفاق وتكرارات الاختلاف بين المتغيرين في الجـدول المـزدوج، ولتوضيح ذلك، فإن الخلية الأولى (ولنرمز لها بالخلية أ) تتضمن 28 تكرارًا، أما الخلية الثانية (ولنرمز لهـا بالخلية ب، فإنها تتضمن 41 تكرارًا، في حين تتضمن الخلية الثالثـة (الخليـة ج) 12 تكـرارًا، وأخـيرًا فـإن الخلية الرابعة (الخلية د) تتضمن 23 تكرارًا، ونتبين من الجدول أن هناك 28 مبحوثًا مـن ذوي المعـدل المنخفض في مشاهدة التليفزيـون ومعـدل قـراءة الصـحف (مـنخفض & مـنخفض)، كـما أن هنـاك 41 مبحوثًا من ذوي المعدل المنخفض في مشاهدة التليفزيون وفي الوقت نفسه معدل متوسـط مـن حيـث قراءة الصحف (منخفض & متوسط)، وهناك 12 مبحوثًا من ذوي المعدل المـنخفض فـيما يخـص قـراءة الصحف ومعدل متوسط من حيث مشـاهدة التليفزيـون (مـنخفض، متوسـط)، أخـيرًا، فـإن هنـاك 23 مبحوثًا ذوي معدل متوسط من حيث مشاهدة التليفزيون وفي الوقت نفسه معدل متوسط من حيـث قراءة الصحف (متوسط & متوسط).

فإذا رجعنا إلى معادلة معامل جاما، فإن الرمز $ن_{ق}$ هو حاصل ضرب عدد حـالات الاتفـاق، ونحصـل عليـه بضــرب تكـرارات الخليـة (ا) في تكـرارات الخليـة (د)؛ أي

28× 23= 644، أما الرمز $ن_{خ}$ فهو حاصل ضرب عدد حالات الاختلاف، ونحصل عليه بضرب تكرارات الخلية (ب) في تكرارات الخلية (ج)؛ أي 41× 12= 492 وتكون قيمة جاما تساوي:

$$ج = \frac{٦٤٤ - ٤٩٢}{٦٤٤ + ٤٩٢} = 0.134$$

غير أنه من الجدير بالذكر أن معامل (جاما) يستخدم أيضًا في التصنيفات الأكثر من ثنائية للمتغيرات، كأن نريد إيجاد العلاقة بين معدل مشاهدة البرامج السياسية في التليفزيون (منخفض، متوسط، مرتفع) ومستوى المعرفة بالقضايا الدولية (منخفض، متوسط، مرتفع)، كما يمكن استخدام معامل (جاما) لحساب الارتباط بين متغير اسمي ومتغير فئوي.

(5) **معامل ارتباط كندال:**

يستخدم معامل ارتباط كندال (Kendal Correlation Coefficient) في قياس الارتباط بين متغيرين كلاهما من النوع الرتبي ويعتمد على نفس فكرة معامل جاما، ويرمز لمعامل ارتباط كندال بالرمز Ta (تو أ) وتقرأ (تو ألف) والقانون المستخدم لذلك على الصورة:

$$تو\ أ = \frac{ن_{ق} - ن_{خ}}{٠,٥\ ن\ (ن - ١)}$$

حيث: تو أ معامل ارتباط كندال من النوع أ

كذلك $ن_{ق}$، $ن_{خ}$ كما هو في حالة معامل جاما (بمعنى أن: ($ن_{ق}$) حاصل ضرب عدد حالات الاتفاق، بينما ($ن_{خ}$) هي حاصل ضرب عدد حالات الاختلاف. أما الحرف (ن) فهو عدد أفراد عينة الدراسة.

فإذا كانت $ن_{ق}$ = 45 بينما $ن_{خ}$ = 23 وكان عدد مفردات العينـة (10) فـإن معامـل ارتبـاط كنـدال يساوي:

$$تو\ أ = \frac{٤٥ - ٢٣}{٠{,}٥ \times ١٠ \times ٩} = \frac{٢٢}{٤٥} = 0.49$$

وفي بعض الحالات، فإن الرتب تكون متكررة، بمعنى أن الكثير من المبحوثين يأخذون الترتيب نفسه على المتغيرين، وهنا لا تصل قيمة معامل ارتباط كنـدال إلى الواحـد الصحيح (الارتبـاط التـام) رغـم أن المتغيرين في الحقيقة قد يرتبطان ارتباطًا تامّا.

(6) معامل اتفاق كندال:

يستخدم معامـل اتفـاق كنـدال (W) (Kendall Coefficient of Concordance) لحسـاب معامـل الاتفاق بين الرتب، فقد يقتضي البحث حساب الارتباط بين أكثر من ترتيبين (وليس ترتيبان اثنـان فقـط كما في حالة ارتباط سبيرمان)، مـن ذلـك مـثلًا تقصـي الارتبـاط بـين: ترتيـب قـراءة الصحف، ترتيـب مشاهدة التليفزيون، ترتيب استخدام الإنترنت، ترتيب الاستماع للراديو. في هذه الحالة نستخدم معامل اتفاق كندال، كما يشيع استخدام هذا المعامل في التحقق من ثبات الاستبيانات وغيرها حين تعرض على مجموعة من المحكمين. لنفرض أننا عرضنا استبانة تتضمن عددًا من البنود على مجموعة من المحكمين؛ بهدف الكشف عن ترتيبهم لهذه البنود من حيث الأهمية والدلالة للبحث، في هذه الحالة، فإنه لمعرفة مدى اتفاق هؤلاء المحكمين في ترتيب البنود يتم استخدام معامل اتفاق كندال، لأنه يسـهل الإجـراءات، ويناسب التعرف على درجة الاتفاق بين الرتب، ويتمثل قانون اتفاق كندال في الصورة الآتية:

$$ر = \frac{٢}{م^{٢}} \times \frac{٦\ مج\ ف^{٢}}{ن\ (ن^{٢} - ١)}$$

حيث م: عدد المحكمين

ن: حجم العينة (عدد المحكوم عليهم بنودًا كانت أم أفرادًا...)

ف^2: مربعات الفروق بين مجموع رتب كل مفردة والمتوسط العام لمجموع الرتب

مثال:

في سياق التخطيط السياسي للتعامل مع قضية الملف النووي الإيراني أعدت وكالة المخابرات المركزية الأمريكية (CIA) إستراتيجية دعائية وإعلامية تتضمن 5 خطط متكاملة لتنفيذها على المستوى المحلي والدولي حيال هذه المسألة، وتم عرض هذه الإستراتيجية على ثلاثة خبراء. ولإقرار هذه الإستراتيجية لابد أن يكون معامل الاتفاق بين هؤلاء الخبراء لا يقل عن 90% بافتراض أن مج ف^2 يساوي 72 فهل سيتم إقرار هذه الإستراتيجية؟

في هذا المثال نلاحظ أن (م) = 3 بينما قيمة ن = (5) أما (مج ف^2) فهو (72)، وبالتالي فإن معامل اتفاق كندال يساوي:

$$\text{ر} = \frac{\text{٢}}{\text{٩}} \times \frac{\text{٤٣٢}}{\text{٥}(\text{٢٤})} = 0.792$$

أي أن نسبة الاتفاق بين الخبراء حول ترتيب أهمية عناصر الإستراتيجية الدعائية الأمريكية بلغت 79%؛ أي أنها تقل عن 90% كنسبة اتفاق مطلوبة بين الخبراء، وبالتالي لن يتم إقرار الإستراتيجية المذكورة.

وكثيرًا ما يستخدم معامل اتفاق كندال في بحوث الاتصال ونحن بصدد التعرف على العلاقة بين ترتيب المفحوصين للبرامج والمواد الإعلامية المتعددة؛ من حيث أهميتها بالنسبة لهم باعتبار هؤلاء المفحوصين من مستخدمي وسائل الإعلام، كما يطبق معامل اتفاق كندال على استجابات القائمين بالاتصال ونحن بصدد رصد وتحليل آرائهم وتقييماتهم بشأن قضايا معينة تتخذ رتبًا من حيث الأهمية... إلخ، هذا بالإضافة إلى استخدام معامل اتفاق كندال في التحقق من صدق أدوات جمع البيانات.

(7) معامل ارتباط فاي:

يستخدم معامل ارتباط فاي (Phi Coefficient) لقياس لعلاقة بين متغيرين اسميين Two Nominal (Variables) ينقسم كل منهما انقسامًا ثنائيًا مثل: (نعم & لا)، (ذكر & أنثى) (موافق & رافض)، (حضر& ريف)... إلخ ويرمز له بالرمز اللاتيني ø. لنفرض مثلًا أن المتغير الأول هو مشاهدة التليفزيون يوميًا (س)، أما المتغير الثاني، فهو قراءة الصحيفة بانتظام (ص)، وهنا يكون توزيع المبحوثين حسب هذين المتغيرين كما يلي:

المشاهدة (س)	القراءة (ص)	
	نعم	**لا**
نعم	أ	ب
لا	ج	د

فالرموز أ، ب، ج، د هي المشاهدات في صورة تكرارات تتوزع على الأقسام المختلفة لكل من مشاهدة التليفزيون يوميًا (س)، وقراءة الصحيفة بانتظام (ص)، ويتم حساب ارتباط فاي (Phi) بين استجابات المفحوصين على هذين المتغيرين باستخدام القانون:

$$\phi = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{ج}}{\sqrt{(\text{أ} + \text{ب})\ (\text{ج} + \text{د})\ (\text{أ} + \text{ج})\ (\text{ب} + \text{د})}}$$

حيث:

ϕ معامل فاي

أ = عدد الأفراد الذين أجابوا (نعم) على السؤال س، وأجابوا (نعم) على السؤال ص

ب = عدد الأفراد الذين أجابوا (نعم) على السؤال س، وأجابوا (لا) على السؤال ص

ج = عدد الأفراد الذين أجابوا (لا) على السؤال س، وأجابوا (نعم) على السؤال ص

د = عدد الأفراد الذين أجابوا (لا) على السؤال س، وأجابوا (لا) على السؤال ص

مثال: الجدول الآتي يوضح الاستجابة على سؤالين، الأول يتعلق بقراءة إحدى الصحف اليومية بانتظام (س)، والثاني يتعلق بقراءة إحدى المجلات الفنية الأسبوعية بانتظام (ص):

ص	س	
	نعم	**لا**
نعم	8	7
لا	5	10

والمطلوب حساب ارتباط (Phi) بين قراءة الصحيفة وقراءة المجلة. في هذا المثال نجد أن:

قيمة (أ) = 8

قيمة (ب) = 7

قيمة (ج)= 5

قيمة (د) = 10

$$\phi = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{ج}}{\sqrt{(\text{أ} + \text{ب})\ (\text{ج} + \text{د})\ (\text{أ} + \text{ج})\ (\text{ب} + \text{د})}}$$

$$= \frac{\text{٨} \times \text{١٠} - \text{٧} \times \text{٥}}{\sqrt{(\text{٨}+\text{٧})(\text{٥}+\text{١٠})(\text{٨}+\text{٥})(\text{٧}+\text{١٠})}}$$

$$= \frac{\text{٨٠} - \text{٣٥}}{\sqrt{(\text{١٥})(\text{١٥})(\text{١٣})(\text{١٧})}}$$

$$= \frac{\text{٤٥}}{\sqrt{\text{٤٩٧٢٥}}}$$

$$= \frac{٤٥}{٢٢٣} = 0.202$$

أي أن هناك ارتباطًا موجبًا قدره 0.202 تقريبًا بين قراءة الصحيفة اليومية بانتظام وقراءة المجلة الفنية بانتظام.

(8) معامل التوافق:

يشيع استخدام معامل التوافق (Contingency Coefficient) في الكشف عن العلاقة أو الارتباط بين متغيرات اسمية أو نوعية (Nominal) مثل متغير الحالة الاجتماعية (متزوج – أعزب – مطلق) وكذلك متغير منطقة الإقامة (منطقة حضرية- منطقة ريفية- منطقة صحراوية)... إلخ، كما يمكن استخدامه أيضًا لقياس الارتباط بين متغير نوعي والآخر رتبي (Ordinal)ويمكن استخدام معامل التوافق، سواء كان الجدول المزدوج من النمط 2×2 أو أكثر، وسواء تساوى عدد خلايا الصفوف وعدد خلايا الأعمدة أو لم يتساو، والقانون المستخدم للحصول على قيمة معامل التوافق هو:

$$ق = \sqrt{١ - \frac{١}{ج}}$$

وإذا كانت (ق) تعني معامل التوافق، فإن (ج) تعني مربع تكرار كل خلية مقسومًا على مجموع تكرارات الصف × مجموع تكرارات العمود، ويتم ذلك لكل خلية من خلايا الجدول؛ أي أن قيمة (ج) لكل خلية نحصل عليها بموجب القانون:

$$ج = \frac{\text{مربع تكرار الخلية}}{\text{مجموع التكرارات لعمود تلك الخلية} \times \text{مجموع التكرارات لصف نفس الخلية}}$$

مثال: الجدول الآتي يوضح توزيع عينة من المفحوصين حسب منطقة الإقامة والصحف التي يقرؤونها:

منطقة الإقامة	الصحف المقروءة			المجموع
	الصحف القومية	الصحف الحزبية	الصحف المستقلة	
القاهرة	4	6	3	13
وجه بحري	3	2	4	9
وجه قبلي	5	7	6	18
المجموع	12	15	13	40

باستبعاد خلايا (المجموع)، فإن الجدول يتضمن (تسع) خلايا، وللحصول على معامل التوافق لا بد من الحصول على القيمة (ج) وذلك لكل خلية من الخلايا التسع، ثم جمع قيم (ج) لكل هذه الخلايا، حتى يمكن تطبيق المعادلة السابقة.

ففيما يخص الخلية الأولى نجد أنها تتضمن 4 تكرارات، ومجموع تكرارات العمود هو 12، كما أن مجموع تكرارات الصف هو 13، وبالتالي فإن (ج) لهذه الخلية يكون:

$$0.103 = \frac{٤ \times ٤}{١٣ \times ١٢}$$

فيما يخص الخلية الثانية نجد أنها تتضمن 6 تكرارات، ومجموع تكرارات العمود هو 15 كما أن مجموع تكرارات الصف هو 13، وبالتالي فإن (ج) لهذه الخلية يكون:

$$0.185 = \frac{٦ \times ٦}{١٣ \times ١٥}$$

وهكذا فيما يخص كل خلية من الخلايا التي تتضمن التكرارات موزعة على المتغيرين محل الدراسة إلى أن نصل إلى الخلية التاسعة، وهي تتضمن 6 تكرارات، ومجموع تكرارات

العمود هو 13، كما أن مجموع تكرارات الصف هو 18، وبالتالي فإن (ج) للخلية التاسعة يكون:

$$\frac{٦ \times ٦}{١٨ \times ١٣} = 0.154$$

وبعد الحصول على قيمة (ج) لكل الخلايا، نقوم بجمع القيم التي حصلنا عليها، لنفرض أن هذه القيم:

1.054= 0.154 +0.181 +0.116 +0.137 +0.03 +0.083 +0.065 + 0.185 +0.103

أي أن القيمة (ج) تساوي 1.054 فنحصل بذلك على معامل التوافق كالآتي:

$$ق = \sqrt{١ - \frac{١}{١,٠٥٤}} = 0.226$$

وعند تفسير معامل التوافق، فإنه يجب تذكر أن قيمته دائمًا موجبة، كما أن القيمة العظمى لمعامل التوافق لا تصل إلى الواحد الصحيح؛ لأن هذه القيمة تعتمد على العدد الأصغر للأعمدة أو الصفوف.

ويتم التعرف على الدلالة الإحصائية لمعامل التوافق من خلال حساب قيمة $كا^2$، وذلك بموجب المعادلة:

$$كا^2 = \frac{ن \times ق^٢}{١ - ق^٢}$$

علمًا بأن الرمز (ن) هو عدد أفراد العينة، أما الرمز ($ق^2$) فهو مربع معامل التوافق المحسوب. وفي مثالنا السابق فإن عدد مفردات العينة هو 40 أما معامل التوافق، فهو 0.226 وعلى ذلك تكون قيمة $كا^2$ تساوي:

$$2.15 = \frac{٠,٠٥١ \times ٤٠}{١ - ٠,٠٥١}$$

أي أن قيمة كا^2 تساوي 2.15 وعلينا أن نقارن هـذه القيمـة المحسـوبة بقيمـة كـا^2 الجدوليـة عنـد درجة حرية تساوي (عـدد الأعمـدة – 1) × (عـدد الصـفوف – 1)، علـمًا بـأن درجـة الحريـة في مثالنـا المذكور هي (4)، وذلك لأن عدد الأعمـدة هـو (3)، وعـدد الصـفوف أيضًـا (3)، وبالتـالي تكـون درجـة الحرية = (3-1) × (3-1) = 4 وعلى الباحث أن يقارن قيمة كـا^2 المحسـوبة (وهـي 2.15) بقيمـة كـا^2 الجدولية عند مستوى ثقة معين (0.95 أو 0.99) فإذا كانت قيمـة كـا^2 المحسـوبة أقـل مـن قيمـة كـا^2 الجدولية، فإن ذلك يعني أنه لا يوجد ارتباط بين الصحف المقروءة ومنطقة الإقامة (بمعنى أن معامـل التوافق غير دال)، أما إذا كانت قيمة كا^2 المحسوبة أكبر من قيمة كا^2 الجدولية، فإن ذلك يعنـي وجـود ارتباط بين الصحف المقروءة ومنطقة الإقامة (بمعنى أن معامل التوافق دال).

وبالرجوع إلى جدول القيم الحرجة لمعامل (كا^2) نتبين أنه عنـد درجـة حريـة (4)، فـإن قيمـة كـا^2 تكون دالة إحصائيًا عند مسـتوى (0.05) إذا كانـت هـذه القيمـة تسـاوي 9.49 ونظـرًا لأن قيمـة كـا^2 المحسوبة تساوي 2.15، فإنه لا يوجد ارتباط بين منطقة الإقامـة والصـحف المقـروءة (إذ إن قيمـة كـا^2 المحسوبة أقل من قيمة كا^2 الجدولية). مرة أخرى، فإن البرامج الإحصائية الجاهزة تمكننا مـن الحصـول على هذه المعطيات الإحصائية بسهولة، لكن المعرفة بها أمر أساسي للبحث العلمي.

(9) اختبار كا^2 للاستقلالية:

يستخدم كا^2 (كا تربيع) في الكثير من التطبيقات، من بينها اختبار كـا^2 للاسـتقلالية (Chi Square Test of Independence)، وفي هذه الحالة يـتم التطبيـق عـلى مجموعـات عينة واحدة لمعرفة مدى اختلافهم أو عدم اختلافهم في جانب معـين (رأي أو صـفة أو سلوك أو موقف أو اتجاه... إلخ)، فالهدف هنا معرفة مدى استقلالية هذا الجانب عن

متغير معين ضمن خصائص العينة.إن مدى الاستقلالية يعني الاختلاف (بمعنى عدم الارتباط)، فقد يريد الباحث- على سبيل المثال- معرفة ما إذا كانت آراء العينة بشأن جدوى استخدام التليفزيون في تنظيم الأسرة (موافق أو غير موافق) لا تختلف باختلاف خصائص هؤلاء المفحوصين؛ من حيث مكان الإقامة (منطقة ريفية، منطقة حضرية، منطقة صحراوية). في هذه الحالة، يتم استخدام اختبار $كا^2$ للاستقلالية، وقانونه هو:

$$كا^2 = مج \frac{(هـ - ق)^٢}{ق}$$

حيث هـ: تعني التكرارات المشاهدة (الموجودة في كل خلية من خلايا الجدول)

ق: تعني التكرارات المتوقعة (التكرارات التي نتوقع أن توجد في كل خلية من خلايا الجدول)، ويتم حساب (ق)؛ أي التكرارات المتوقعة بموجب القانون:

$$ق = \frac{\text{مجموع تكرار الصف × مجموع تكرارات العمود}}{\text{عدد العينة}}$$

ويتم تطبيق هذا القانون لكل خلية من خلايا الجدول الذي يوضح توزيع العينة حسب المتغيرين المراد تقصي العلاقة بينهما. إن الباحث قد يريد التحقق من صحة الفرض القائل: «لا تختلف نسبة الموافقين على جدوى استخدام التليفزيون في تنظيم الأسرة باختلاف منطقة الإقامة».

مثال: الجدول الآتي يوضح توزيع عينة من المبحوثين (ن= 200) حسب منطقة الإقامة والموافقة على استخدام التليفزيون في حملات تنظيم الأسرة:

منطقة الإقامة	الموافقة		المجموع
	موافق	غير موافق	
منطقة ريفية	42	18	60
منطقة حضرية	33	23	56
منطقة صحراوية	45	39	84
المجموع	**120**	**80**	**200**

إذا استبعدنا خلايا المجموع، فإن هذا الجدول يتضمن 6 خلايا، وتتضمن كل خلية عددًا من التكرارات. وبموجب القانون السابق ذكره، يتم حساب التكرار المتوقع لكل خلية، على سبيل المثال، فإن الخلية الأولى (التي تتضمن 42 تكرارًا) نلاحظ من الجدول أن مجموع تكرارات عمود تلك الخلية هو (120)، أما مجموع تكرارات صف تلك الخلية فهو (60)، وبالتالي يكون التكرار المتوقع للخلية الأولى:

(ق) الخلية الأولى $\frac{٦٠ \times ١٢٠}{٢٠٠} = 36$

وهكذا بالنسبة لجميع خلايا الجدول، إلى أن نصل إلى الخلية السادسة والأخيرة؛ حيث نلاحظ أن مجموع تكرارات عمود تلك الخلية هو (80) أما مجموع تكرارات صف تلك الخلية فهو (84)، وبالتالي يكون التكرار المتوقع للخلية السادسة:

(ق) الخلية السادسة $\frac{٨٤ \times ٨٠}{٢٠٠} = 33.6$

بهذه الطريقة يتم حساب التكرار المتوقع لكل خلية من خلايا الجدول، فيكون لدينا ستة تكرارات متوقعة (بواقع تكرار لكل خلية)، بعد ذلك نستخدم التكرارات الموجودة بالجدول والتكرارات المتوقعة (التي تم حسابها) لمعرفة قيمة كا^2 لكل خلية من خلايا الجدول. كمثال على ذلك فإن التكرارات الموجودة (المشاهدة) في الخلية الأولى هي 42 تكرارًا، أما التكرارات المتوقعة التي تم حسابها لهذه الخلية فتبلغ 36، ويتم حساب قيمة كا^2 للخلية الأولى كالآتي:

$$\frac{(٤٢ - ٣٦)^٢}{٣٦} = 1$$

وهكذا بالنسبة لجميع الخلايا إلى أن نصل إلى الخلية السادسة؛ حيث نلاحظ من الجـدول أن هـذه الخلية تتضمن (39) تكرارًا، أمـا التكـرارات المتوقعـة التـي تـم حسـابها للخليـة السادسـة فتبلـغ 33.6 وبالتالي فإن قيمة كا^2 للخلية السادسة تساوي:

$$0.87 = \frac{(٣٩ - ٣٣,٦)^٢}{٣٣,٦}$$

وبهذا نحصل على قيمة كا^2 لكل خلية من خلايا الجدول، فيكون لدينا ست قيم. إن مجمل هذه القـيم الست هو كا^2 للجدول ككل، لنفرض أن هذه القيم هي: (1)، (1.5)، (0.01)، (0.02) (0.58)، (0.87)

إن مجموع هذه القيم يساوي (3.98)، وعلينا أن نقارن هـذه القيمـة بقيمـة كـا^2 الجدوليـة عنـد درجة حرية تساوي:

(عدد الأعمدة – 1) (عدد الصفوف – 1)

وفي مثالنا المذكور، فإن الجدول السابق يوضح أن عدد الأعمدة هو 2 (موافق & غير موافق) بيـنما عدد الصفوف هو 3 (منطقة ريفية، منطقة حضـرية، منطقة صحراوية)، وبالتالي فإن درجة الحريـة = 2= (1-3) (1-2)

وعند درجة حرية 2 تتم مقارنة قيمة كا^2 المحسوبة بقيمة كا^2 الجدولية، فإذا كانت قيمـة كـا^2 المحسـوبة أقل من قيمة كا^2 الجدولية، فإننـا نقبـل الفـرض الصـفري (لا تختلـف نسـبة المـوافقين عـلى جـدوي اسـتخدام التليفزيون في تنظيم الأسرة باختلاف منطقة الإقامة)،أما إذا كانت قيمة كا^2 المحسوبة أكبر من أو تساوي قيمـة كا^2 الجدولية، فإننا نرفض الفرض الصفري ونقبل الفرض البديل (تختلف نسبة الموافقين عـلى جـدوى اسـتخدام التليفزيـون في تنظـيم الأسرة بـاختلاف منطقـة الإقامـة). وبـالرجوع إلى الجـداول الإحصـائية (جـدول كـا^2)

نجد أنه عند درجة حرية 2 فإن قيمة كا^2 تساوي (5.99) عند مستوى الثقة 0.95 أما عند مستوى الثقة 0.99 فإن قيمة كا^2 تساوي (9.21)، ولما كانت قيمة كا^2 المحسوبة تساوي (3.98)، فإنها بذلك تكون أقل من قيمة كا^2 الجدولية، وبالتالي نقبل الفرض الصفري القائل بعدم وجود اختلاف أو علاقة بين المتغيرين؛ أي أن الموافقة على استخدام التليفزيون في حملات تنظيم الأسرة لا تختلف باختلاف منطقة الإقامة. لا شك في أن البرامج الإحصائية الجاهزة تتيح لنا هذه المعطيات، غير أنه من الضروري فهم هذه الأسس.

(10) معامل ثيتا (معامل فريمان)

يستخدم معامل ثيتا (Theta Coefficient) أو معامل فريمان (Freeman Coefficient) إذا كانت البيانات تتعلق بمتغيرين أحدهما اسمي، الجنس مثلًا (ذكور & إناث) أو الجنسية (مصري & غير مصري) والمتغير الآخر في صورة رتبة (أشاهد التليفزيون يوميّا، أشاهد التليفزيون معظم أيام الأسبوع، أشاهد التليفزيون بعض أيام الأسبوع، أشاهد التليفزيون يوم واحد في الأسبوع، لا أشاهد التليفزيون) ولهذه الاستجابات أرقام رتبية (5، 4، 3، 2، 1) ويجب أن تكون البيانات عند استخدام هذا المعامل موضحة تميز تكرارات المتغير الاسمي في كل مستوى رتبي، بمعنى أننا إذا كنا بصدد بحث العلاقة بين الجنس (ذكور & إناث) ومشاهدة التليفزيون وفق الرتب المذكورة، فإنه يجب تميز الذكور على الإناث أو العكس في كل مستوى رتبي، كأن يتميز الذكور عن الإناث في المستوي الرتبي الأول، وتتميز الإناث عن الذكور في المستوى الرتبي الرابع... وهكذا في بقية المستويات الرتبية، والتميز هنا يقصد به عدد التكرارات في المستويات الرتبية المختلفة، فإذا كان عدد الإناث أكبر من عدد الذكور في مستوى رتبي معين، فإن ذلك يعني تميز الإناث على الذكور في هذا المستوى، ولا يشترط أن تكون البيانات الرتبية متسلسلة، ويرمز لمعامل ثيتا بالرمز اللاتيني Ǿ، وكمثال على هذا المعامل

نفرض أننا بصدد بحث معامل الارتباط بين النوع (ذكور & إناث) وترتيب الـدرجات في مـادة الإحصـاء (1، 2، 3، 4، 5) لكل مفردة من الجنسين، في هذه الحالة، فإن:

$$\text{ثيتا} = \frac{\text{ك ذ ث} - \text{ك ث ذ}}{\text{ن ذ} \times \text{ن ث}}$$

حيث إن:

- ك ذ ث: عدد التكرارات التي يكون فيها الذكور في رتبة أعلى من الإناث
- ك ث ذ: عدد التكرارات التي يكون فيها الإناث في رتبة أعلى من الذكور
- ن ذ: عدد مجموعة الذكور
- ن ث: عدد مجموعة الإناث

كمثال توضيحي نفرض أن دراسة أجريت على عينة قوامها 200 مفـردة مـن الجنسـين (160 ذكـور، 40 إناث)، وعند تنظيم البيانات تبين أن استجابات الإناث عكست أنهن أتـين بواقـع 30 تكـرارًا في رتبـة أعلى من الذكور؛ من حيث معدل مشاهدة برامج الشباب في التليفزيـون، في حـين جـاء الـذكور بواقـع 120 تكرارًا في مرتبة أعلى من الإناث من حيث المعدل نفسه. في هذه الحالة فإن (ك ث ذ) تسـاوي 30 أما (ك ذ ث) فتساوي 120، وبالتالي فإن معامل ثيتا يكون:

$$\text{ثيتا} = \frac{١٢٠ - ٣٠}{١٦٠ \times ٤٠} = 0.014$$

أي أن معامل ثيتا يساوي 0.014 وهو ارتباط ضعيف كما هو واضح.

تفسير معامل الارتباط

يتضح مما سبق أن معامل الارتباط بين متغيرين هـو قيمـة مجـردة تعـبر عـن العلاقـة القائمـة بـين المتغيرين؛ بحيث تنحصر بين (+1)، (-1)، وفيما بين هـاتين القيمتـين يعـبر عـن قيمـة معامـل الارتبـاط بكسر عشري.

وهناك أخطاء شائعة في تفسير معامل الارتباط، فقد يتعامل الباحث مع قيمة معامل الارتباط على أنها قيمة مطلقة مثل قيمة الطول أو الوزن مثلًا، أو على أنه نسبة مئوية وهذا غير صحيح، فمثلًا معامل الارتباط (0.25) لا يعد نصف معامل الارتباط (0.50)، ومعامل الارتباط (0.50) لا يعد نصف معامل الارتباط الذي قيمته واحد صحيح،كما أن الفرق بين معاملي الارتباط (0.40)، (0.60) لا يساوي الفرق بين معاملي الارتباط (0.70)، (0.90)

إن معامل الارتباط هو مقدار مجرد ولا يقاس على ميزان خطي وحداته متساوية، كما لا يجب تفسير معامل الارتباط على أساس وحدات الدرجات الأصلية، فقد يكون اهتمام الصحيفة بقضايا المرأة مقاسًا بمساحة المادة الصحفية التي استغرقتها هذه القضايا (المساحة بالسنتيمتر المربع)، أما اهتمام الراديو بقضايا المرأة، فقد يكون مقاسًا بالوقت أو المدة الزمنية؛ أي أن وحدة القياس مختلفة في الحالتين، وعند إيجاد معامل الارتباط بين اهتمام الصحيفة بقضايا المرأة (المتغير الأول) واهتمام الراديو بهذه القضايا (المتغير الثاني) فإن قيمة معامل الارتباط تكون مستقلة عن الوحدات التي يقاس بها المتغيران والقيم التي يأخذها كل منهما.

كما أن بعض الباحثين يعتبر أن معامل الارتباط الذي تنحصر قيمته بين 0.30، 0.70 هو ارتباط متوسط القيمة؛ أي أنه يعبر عن علاقة ارتباطية متوسطة، بينما يعتبر أن معامل الارتباط الذي تقل قيمته عن ذلك منخفضًا، أما إذا زادت قيمته عن 0.70، فإنه يعتبر مرتفعًا، لكن هذه الاعتبارات غير صحيحة، ذلك أن دلالة معامل الارتباط هي دالة لحجم العينة، فقد يكون معامل ارتباط قدره (0.3)، ولكنه دال إحصائيًا (Significant) لأن حجم العينة 600 مفردة مثلًا، وفي الوقت نفسه فإن معامل ارتباط قدره (0.7) لكنه غير دال إحصائيًا (Not Significant) لأن حجم العينة عشر مفردات؛ أي أن قيمة

معامل الارتباط المرتفعة التي يحصل عليها الباحث باستخدام عينات صغيرة قد لا يكون لها أي دلالة إحصائية.

من جهة أخرى فإن مقدار العلاقة بين متغيرين لا تعتمد على إشارة معامل الارتباط، فمعامل الارتباط (-0.70) يعبر عن نفس مقدار العلاقة بين متغيرين معامل الارتباط بينهما (+0.70) فالفرق بينهما يكون في اتجاه العلاقة وليس في قيمة العلاقة.

أخيرًا، فإن قيمة معامل الارتباط لا تتأثر بإضافة أو طرح مقدار ثابت إلى أو من كل قيمة من قيم أحد المتغيرين، فقد يجد الباحث مثلًا أن قيمة معامل الارتباط بين معدل استخدام التليفزيون ومعدل استخدام الإنترنت يساوي (0.6) وفي تطبيق آخر على المجموعة نفسها يتضح أن معامل الارتباط بنفس القيمة (0.6)، رغم أن متوسط درجات المجموعة على الاختبارين في التطبيق الثاني مختلف عن متوسط درجتها في التطبيق الأول، وعلى الرغم من ذلك فإن معامل الارتباط ظل كما هو.

إن ذلك يفسر بما سبق ذكره وهي أن قيمة معامل الارتباط لا تتأثر بإضافة أو طرح مقدار ثابت إلى أو من كل قيمة من قيم أحد المتغيرين، ففي التطبيق الثاني قد يكون متوسط درجة المجموعة فيما يخص معدل استخدام التليفزيون قد زاد بقيمة معيارية تساوي الزيادة في معدل استخدام الإنترنت (لقد زاد كلا المتغيرين بقيمة ثابتة أو واحدة)، وبالمثل فقد يكون المتغيران قد نقصا بقيمة واحدة، فمعامل الارتباط هو قيمة تدل على التغاير أو التباين المتلازم (Concomitant Variation) بين المتغيرين، ولا يشير إلى مقدار المتغيرين.

ومن الطرق المفيدة في تفسير القيم المختلفة لمعامل الارتباط (ر) هو تربيع هذه القيم؛ أي الحصول على قيمة (ر^2)، والمقدار (ر^2) هو النسبة بين التباين الكلي لأحد المتغيرين والجزء من هذا التباين الذي يمكن التنبؤ به باستخدام المتغير الثاني. أي أن (ر^2) هي الجزء من تباين أحد المتغيرين الذي يمكن أن نتنبأ به باستخدام المتغير الثاني. فإذا كان معامل

الارتباط بين متغيرين هو(0.707) مثلًا فإن ر^2 = (707)^2 =0.5 تقريبا، وعندما تكون قيمة ر= 0.50 فإن ر^2 = 0.25؛ ولذلك فإنه يمكن اعتبار أن معامل الارتباط 0.707 ضعف معامل الارتباط 0.50 حيث إن نسبة(ر^2) في الحالتين هي (1:2) تقريبا

رابعًا: الارتباط المتعدد

معامل الارتباط المتعدد (R) هو الارتباط بين متغير تابع وأكثر من متغير مستقل، من ذلك مثلًا الارتباط بين الدرجة على مقياس المعرفة السياسية (كمتغير تابع) والدرجة على مقياس استخدام كل من: قراءة الصحف اليومية، مشاهدة المواد الإخبارية في التليفزيون، التفاعل مع المنتديات السياسية عبر الإنترنت (كمتغيرات مستقلة)، وتتراوح قيمة معامل الارتباط المتعدد بين الصفر والواحد الصحيح الموجب؛ أي أنه موجب القيمة دائمًا، وهذه القيمة في حقيقتها هي الجذر التربيعي لمعامل التحديد (Determination Coefficient) ورمزه (R^2)، فهو مربع معامل الارتباط، ويدل معامل التحديد على نسبة التباين في المتغير التابع بفعل المتغيرات المستقلة، فإذا كان معامل التحديد R^2 للعلاقة بين متغير تابع ومجموعة متغيرات مستقلة هو 0.68، فإن ذلك يعني أن هذه المتغيرات المستقلة تفسر 0.68 من التباين في المتغير التابع، أما النسبة الباقية (0.32)، فإنها تعزى إلى عوامل أخرى. وإذا كان معامل الارتباط المتعدد موجب القيمة دائمًا، فإن هذه القيمة تكون أكبر من قيمة معامل الارتباط بين أي متغيرين من المتغيرات التي يتم تقصي العلاقة بينها.

وتتعدد صيغ الحصول على معامل الارتباط المتعدد ولكل منها أسسها الجبرية والرياضية، ومن هذه الطرق: صيغة مجموع المربعات، صيغة معاملات الانحدار المعيارية، صيغة معامل ارتباط بيرسون، صيغة الارتباطات الثنائية بين المتغيرات، وهذه الصيغة الأخيرة هي الأكثر ملاءمة لبحوث الاتصال، كما أنها أبسط الطرق لأنها تعتمد على قيمة معامل الارتباط بين كل زوجين من المتغيرات محل الدراسة. ولتبسيط فهم معامل

الارتباط المتعدد بموجب هذه الصيغة، نفرض أن لدينا متغيرًا تابعًا ومتغيرين مستقلين، ولنرمز للمتغير التابع بالرقم (1)، أما المتغير المستقل الأول، فنرمز له بالرقم (2)، والمتغير المستقل الثاني نرمز له بالرقم (3). إن معامل الارتباط المتعدد هو الجذر التربيعي لناتج هذه المعادلة:

$$\text{ر}^2 = \frac{\text{ر}^2\,\text{ت س} + \text{ر}^2\,\text{ت ص} - 2\,(\text{ر ت س})\,(\text{ر ت ص})\,(\text{ر س ص})}{1 - \text{ر}^2\,\text{س ص}}$$

علمًا بأن:

ر^2 : تعني معامل التحديد (مربع معامل الارتباط)، وهو في مثالنا هذا يعني نسبة التباين في المتغير التابع بفعل المتغيرين المستقلين

ر^2 ت س: تعني مربع معامل الارتباط بين المتغير التابع والمتغير المستقل الأول

ر^2 ت ص: تعني مربع معامل الارتباط بين المتغير التابع والمتغير المستقل الثاني

ر^2 س ص: تعني مربع معامل الارتباط بين المتغير المستقل الأول والمتغير المستقل الثاني

مثال:

فيما يلي معاملات الارتباط بين المعرفة السياسية (متغير تابع)، ولنرمز له بالرمز (ت)، وكل من قراءة الصحف اليومية بانتظام (متغير مستقل أول)، ولنرمز له بالرمز (س) والمشاركة في المنتديات السياسية عبر الإنترنت (متغير مستقل ثان) ولنرمز له بالرمز (ص)

متغيرات	ت	س	ص
ت	1		
س	0.58	1	
ص	0.72	0.18	1

فإذا أردنا معرفة الارتباط المتعدد بين المعرفة السياسية (متغير تابع)، وكل من قراءة الصحف اليومية بانتظام، والمشاركة في المنتديات السياسية عبر الإنترنت (كمتغيرين

مستقلين)، فإن ذلك يكون بالتعويض في المعادلة السابقة فنحصل على قيمة معامل التحديد، ويكون معامل الارتباط المتعدد هو الجذر التربيعي لهذه القيمة على النحو الآتي:

$$\text{ر}^{2} = \frac{(٠,٥٨)^{٢} + (٠,٧٢)^{٢} - ٢(٠,٥٨)(٠,٧٢)(٠,١٨)}{١-(٠,١٨)^{٢}}$$

$$\text{ر}^{2} = \frac{٠,٨٥٤٨ - ٠,١٥٠٣٣٦}{٠,٩٦٧٦} = 0.73$$

ويكون معامل الارتباط المتعدد هو $\sqrt{٠,٧٣} = 0.85$

أي أن هناك ارتباطًا قيمته (0.85) بين المعرفة السياسية (كمتغير تابع) وكلٍّ من: قراءة الصحف اليومية بانتظام، المشاركة في المنتديات السياسية (كمتغيرات مستقلة)، ومن الواضح أن هذه القيمة موجبة، كما أنها أكبر من قيمة معامل ارتباط بين أي زوجين من المتغيرات المبينة بالجدول.

وفي ضوء قيمة معامل التحديد (R^2)، فإن قراءة الصحف اليومية بانتظام، والمشاركة في المنتديات السياسية عبر الإنترنت يفسران 0.73 من التباين في المعرفة السياسية.

خامسًا: الارتباط الجزئي

يمثل الارتباط الجزئي (Partial Correlation) أحد تطبيقات الضبط الإحصائي المهمة، وجوهر هذا الارتباط قياس العلاقة المستقيمة بين متغيرين بعد عزل تأثير المتغيرات الأخرى. لنفرض أن لدينا المتغيرات الثلاثة الآتية:

المتغير الأول: المعرفة بالأحداث السياسية

المتغير الثاني: معدل استخدام التليفزيون

المتغير الثالث: معدل قراءة الصحف

باستخدام معامل الارتباط الجزئي يمكننا حساب معامل الارتباط بين أي متغيرين اثنين من تلك المتغيرات بعد تثبيت أثر المتغير الثالث؛ بحيث لا يؤثر في قيمة معامل الارتباط. على سبيل المثال يمكن معرفة الارتباط بين المعرفة بالأحداث السياسية ومعدل استخدام التليفزيون بعد تثبيت معدل قراءة الصحف، كما يمكن معرفة الارتباط بين المعرفة بالأحداث السياسية ومعدل قراءة الصحف بعد تثبيت معدل استخدام التليفزيون.... لقد تم عزل تأثير متغير ثالث من العلاقة بين متغيرين.

إن معامل الارتباط الجزئي في هذه الحالة يكون من الرتبة الأولى (First Order Partial Correlation)؛ لأن هذا أبسط مستوى للارتباط الجزئي؛ حيث المطلوب معرفة معامل الارتباط بين متغيرين بعد تثبيت تأثير متغير ثالث في تلك العلاقة؛ أي أن العمليات الإحصائية تتناول ثلاثة متغيرات، مع تثبيت أحدهم ثم بحث العلاقة بين الاثنين الآخرين.

مثال تطبيقي:

في دراسة عن العلاقة بين استخدام الصحف اليومية (متغير أول)، واستخدام الإنترنت (متغير ثان)، وقراءة الكتب (متغير ثالث)، تبين أن الارتباط بين كثافة استخدام الصحف اليومية وكثافة استخدام الإنترنت = 0.68 ولنرمز له بالرمز (ر$_{1،2}$)، وتنطق راء واحد اثنين (بمعنى الارتباط بين المتغير الأول والمتغير الثاني)، أما الارتباط بين كثافة استخدام الصحف اليومية وقراءة الكتب فقد كان = 0.54 ولنرمز له بالرمز (ر$_{1،3}$)، وتنطق راء واحد ثلاثة (بمعنى الارتباط بين المتغير الأول والمتغير الثالث)، أخيراً فإن الارتباط بين كثافة استخدام الانترنت وقراءة الكتب كانت قيمته 0.72 ولنرمز له بالرمز (ر$_{2،3}$)، وتنطق راء اثنين ثلاثة (بمعنى الارتباط بين المتغير الثاني والمتغير الثالث).

فإذا أردنا حساب معامل الارتباط الجزئي بين أي متغيرين مع تثبيت المتغير الآخر، فـإن ذلـك يمكـن على النحو الآتي:

(1) حساب الارتباط بين كثافة استخدام الصحف اليومية وكثافة استخدام الإنترنت بعد عـزل تـأثير قراءة الكتب ($\text{ر}_{1، 2 - 3}$)، ويكون ذلك بموجب المعادلة:

$$\text{ر}_{1،2-3} = \frac{\text{ر}_{١،٢} - (\text{ر}_{١،٣} \times \text{ر}_{٢،٣})}{\sqrt{\left[١ - (\text{ر}_{١،٣})^{٢}\right]\left[١ - (\text{ر}_{٢،٣})^{٢}\right]}}$$

$$\text{ر}_{1،2-3} = \frac{٠,٦٨ - (٠,٥٤ \times ٠,٧٢)}{\sqrt{\left[١ - (٠,٥٤)^{٢}\right]\left[١ - (٠,٧٢)^{٢}\right]}} = 0.498$$

(2) حساب الارتباط بين كثافة استخدام الصحف وقراءة الكتب بعد عزل تأثير استخدام الإنترنت ($\text{ر}_{1، 3 - 2}$)، ويكون ذلك بموجب المعادلة:

$$\text{ر}_{1،3-2} = \frac{\text{ر}_{١،٣} - (\text{ر}_{١،٢} \times \text{ر}_{٢،٣})}{\sqrt{\left[١ - (\text{ر}_{١،٢})^{٢}\right]\left[١ - (\text{ر}_{٢،٣})^{٢}\right]}}$$

$$\text{ر}_{1،3-2} = \frac{٠,٥٤ - (٠,٦٨ \times ٠,٧٢)}{\sqrt{\left[١ - (٠,٦٨)^{٢}\right]\left[١ - (٠,٧٢)^{٢}\right]}} = 0.01$$

(3) حساب الارتباط بين كثافة استخدام الإنترنـت وقـراءة الكتـب بعـد عـزل تـأثير قـراءة الصـحف اليومية ($\text{ر}_{2،3-1}$)، ويكون ذلك بموجب المعادلة:

$$\text{ر}_{2،3-1} = \frac{\text{ر}_{٢،٣} - (\text{ر}_{١،٢} \times \text{ر}_{١،٣})}{\sqrt{\left[١ - (\text{ر}_{١،٢})^{٢}\right]\left[١ - (\text{ر}_{١،٣})^{٢}\right]}}$$

$$ر_{1-3،2} = \frac{٠,٧٢ - (٠,٦٨ \times ٠,٥٤)}{\sqrt{[١ - (٠,٦٨)^{٢}][١ - (٠,٥٤)^{٢}]}} = 0.70$$

لكن الكثير من الدراسات العلمية لا تقتصر على ثلاثة متغيرات، وإنما تشمل متغيرات عديدة، (مثل: المعرفة السياسية، معدل مشاهدة التليفزيون، معدل قراءة الصحف، معدل سماع الراديو، معدل استخدام الإنترنت، مطالعة الكتب المعنية بالشئون العامة، الاتجاه السياسي، المشاركة في مؤسسات المجتمع المدني.... إلخ)، وقد تقتضي الدراسة مثلًا رصد وتحليل العلاقة بين المعرفة السياسية ومعدل قراءة الصحف بعد عزل تأثير المتغيرات الأخرى (معدل مشاهدة التليفزيون، معدل سماع الراديو، معدل استخدام الإنترنت، مطالعة الكتب المعنية بالشئون العامة، الاتجاه السياسي، المشاركة في مؤسسات المجتمع المدني) في هذه الحالة يكون معامل الارتباط الجزئي من الرتبة الثانية (Second Order Partial Correlation).

ومن الطبيعي أن يتم رصد معامل الارتباط (Correlation) بين المتغيرين المعنيين قبل تثبيت المتغير الثالث (أو المتغيرات الأخرى)، وبموجب هذا التثبيت والحصول على قيمة معامل الارتباط الجزئي (Partial Correlation)، بعد التثبيت، نكون أمام أحد الاحتمالات الثلاثة الآتية:

- **الاحتمال الأول:** أن تكون قيمة معامل الارتباط الجزئي أقل من قيمة معامل الارتباط، وهذا يعني أن المتغير المعزول كان يزيد العلاقة بين المتغيرين الآخرين، على سبيل المثال، فإن معامل ارتباط بيرسون بين المعرفة السياسية وقراءة الصحف قد يكون 0.56 وبعد عزل تأثير مشاهدة التليفزيون، يتضح أن قيمة معامل الارتباط الجزئي بين المعرفة السياسية وقراءة الصحف أصبحت 0.43 لقد انخفض الارتباط بين المتغيرين، هذا يعني أن مشاهدة التليفزيون كان لها تأثير موجب في العلاقة بين المعرفة السياسية وقراءة الصحف.

- **الاحتمال الثاني:** أن تكون قيمة معامل الارتباط الجزئي أعلى من قيمة معامل الارتباط، وهذا يعني أن المتغير المعزول كان يضعف العلاقة بين المتغيرين الآخرين (على سبيل المثال، نفرض أن معامل ارتباط بيرسون بين المعرفة السياسية وقراءة الصحف كان 0.56 وبعد عزل تأثير استخدام الإنترنت تبين أن قيمة معامل الارتباط الجزئي بين المعرفة السياسية وقراءة الصحف أصبحت 0.67 لقد ارتفع الارتباط بين المتغيرين بعد تثبيت استخدام الإنترنت، هذا يعني أن استخدام الإنترنت كان له تأثير سالب في العلاقة بين المعرفة السياسية وقراءة الصحف.

- **الاحتمال الثالث:** أن تتساوى قيمة معامل الارتباط الجزئي مع قيمة معامل الارتباط، وهذا يعني أن المتغير المعزول لا تأثير له في العلاقة بين المتغيرين اللذين نتقصى الارتباط بينهما.

تفسير الارتباط الجزئي:

كثيرًا ما يفسر الارتباط الجزئي باستخدام مفهوم التباين المشترك، وقد سبقت الإشارة إلى أنه قبل تطبيق الارتباط الجزئي يكون لدينا معرفة بقيم معاملات الارتباط بين المتغيرات؛ أي أنه من الضروري رصد معامل الارتباط بين المتغير الأول والمتغير الثاني قبل تثبيت المتغير الثالث. لنفرض أن قيمة معامل ارتباط بيرسون بين هذين المتغيرين كانت =0.6471 وعند تطبيق الارتباط الجزئي - أي تثبيت المتغير الثالث وبحث الارتباط بين المتغير الأول والمتغير الثاني- تبين أن قيمة الارتباط الجزئي 0.6896

إن التباين المشترك الناتج عن إسهام المتغير الثالث يمكن الحصول عليه بسهولة من خلال مربع قيمة الارتباط الجزئي مطروحًا منه مربع قيمة ارتباط بيرسون:

أي $(0.6896)^2 - (0.6471)^2$

أي $0.4755 - 0.4187 = 0.06$

بقسمة هذه القيمة (0.06) على مربع معامل ارتباط بيرسون والضرب في 100 نحصل على النسبة المئوية للارتباط الناتج عن إسهام المتغير الثالث، وكما سبقت الإشارة، فإن مربع معامل ارتباط بيرسون (أي قبل تثبيت المتغير الثالث) يبلغ 0.4187 وبالتالي فإن:

$$\%14.3 = 100 \times 0.4187 \div 0.06$$

أي أن النسبة المئوية للارتباط الناتج عن إسهام المتغير الثالث هي 14.3%، وهذا يعني أن النسبة المئوية للارتباط الناتج عن إسهام متغيرات أخرى يساوي 85.7 %

الارتباط الجزئي وارتباط الجزء:

يختلف الارتباط الجزئي (Partial Correlation) عن الارتباط شبه الجزئي Semi Partial Correlation)، أو ما يعرف أحيانًا بارتباط الجزء (Part Correlation)، ذلك أن الارتباط شبه الجزئي أو ارتباط الجزء يعني عزل تأثير متغير معين من (أحد) المتغيرين اللذين نبحث العلاقة بينهما وليس من هذين المتغيرين مجتمعين، ففي مثالنا السابق تم رصد معامل الارتباط الجزئي بين المعرفة بالأحداث السياسية ومعدل استخدام التليفزيون بعد تثبيت معدل استخدام الراديو، لقد تم تثبيت معدل استخدام الراديو بالنسبة للمتغيرين الآخرين (المعرفة بالأحداث السياسية ومعدل استخدام التليفزيون).

وهنا نكون قد طبقنا الارتباط الجزئي، أما في الارتباط شبه الجزئي أو ارتباط الجزء، فإننا نقوم بتثبيت معدل استخدام الراديو عن معدل استخدام التليفزيون فقط، وليس عن المعرفة السياسية، إن ذلك يعني إبقاء تباين متغير المعرفة السياسية كما هو، وقد يرى الباحث أن مبررات ذلك تتمثل في وجود شكوك حول فاعلية الراديو في المعرفة السياسية. ويلعب الارتباط شبه الجزئي، أو ارتباط الجزء دورًا هامّا في الكثير من التحليلات الإحصائية الأخرى مثل الارتباط المتعدد والتحليل العاملي، وكغيره من معاملات الارتباط، فإن الارتباط الجزئي وشبه الجزئي لا يعني علاقة السبب بالنتيجة، فوجود ارتباط دال إحصائيّا لا يعني أن أحد المتغيرين سبب أو نتيجة للمتغير الآخر.

خلاصة الفصل الثالث

تناول هذا الفصل أهم مقاييس العلاقة والارتباط، وتتلخص موضوعاته الأساسية في:

- الارتباط (Correlation) بمعناه الإحصائي هو اقتران التغير في ظاهرة معينة بالتغير في ظاهرة أخرى، بمعنى أنه إذا تغيرت الظاهرة (س) تتغير الظاهرة (ص)، وهذا التغير يعبر عنه بمعامل الارتباط، وفي المتغيرات الكمية، فإن الارتباط من حيث طبيعته قد يكون موجبًا أو سالبًا: أما من حيث الاتجاه (Direction)، فإن الارتباط قد يكون طرديًا أو عكسيًا، فالارتباط الطردي يعني تغير الظاهرتين في اتجاه واحد (بالزيادة أو النقصان)، أما الارتباط العكسي، فيعني تغير الظاهرتين في اتجاهين متناقضين كزيادة الظاهرة (س) عندما تتناقص الظاهرة (ص)، أو تناقص الظاهرة (س) عندما تزيد الظاهرة (ص).

- تتراوح قيمة معامل الارتباط بين + 1 إلى – 1؛ أي من الارتباط الموجب التام إلى الارتباط السالب التام، ويصل الارتباط إلى نهايته العظمى عندما يقترن تغير المتغير الأول اقترانًا تامًا بالتغير في المتغير الثاني، ولا تتأثر قيمة معامل الارتباط بزيادة أو نقصان تكرارات المتغيرين بكميات ثابتة.

- تتعدد معاملات الارتباط، ويتوقف اختيار معامل الارتباط المناسب على نوعية المتغيرات المطلوب حساب الارتباط بينها؛ أي ما إذا كانت تلك المتغيرات اسمية (Nominal) أو رتبية (Ordinal) أو فاصلة (Interval).

- يستخدم ارتباط بيرسون لحساب الارتباط بين متغيرين كليهما من النوع الفتري، وتتعدد صيغ الحصول على معامل ارتباط بيرسون بين متغيرين.

- يستخدم معامل الاقتران (Association Coefficient) في الكشف عن الارتباط بين متغيرين اسميين أو نوعيين (Nominal) كلاهما ثنائي التقسيم مثل النوع (ذكور &إناث)، قراءة الصحف (يقرأ & لا يقرأ).... وهكذا.

- يستخدم معامل ارتباط سبيرمان للرتب (Spearman's Coefficient of Rank Correlation) لقياس التغير الاقتراني القائم بين ترتيب الأفراد أو الأشياء في صفة معينة وترتيبهم في صفة أخرى، كما يمكن استخدام ارتباط سبيرمان إذا كان أحد المتغيرين أو كلاهما من النوع الفئوي أو النسبي، وذلك بعد تحويل البيانات إلى رتب.

- يستخدم ارتباط جاما (Gamma) لمعرفة الارتباط بين المتغيرات الرتبية، ويشيع استخدامه بكثرة إذا كان لدينا رتبتان فقط لكل متغير، خاصة عندما تتكرر بعض القيم في المتغير الواحد.

- يستخدم معامل ارتباط فاي (Phi Coefficient)؛ لقياس لعلاقة بين متغيرين اسميين (Two Nominal Variables) ينقسم كل منهما انقسامًا ثنائيّا مثل: (نعم & لا)، (ذكر & أنثى) (موافق & رافض)، (حضر& ريف)... إلخ، ويرمز له بالرمز اللاتيني ø.

- يستخدم معامل ارتباط كندال (Kendal Correlation Coefficient) في قياس الارتباط بين متغيرين كلاهما من النوع الرتبي، ويعتمد على نفس فكرة معامل جاما، ويرمز لمعامل ارتباط كندال بالرمز Ta (تو أ) وتقرأ (تو ألف).

- يستخدم معامل اتفاق كندال (W) (Kendall Coefficient of Concordance) لحساب معامل الاتفاق بين الرتب، فقد يقتضي البحث حساب الارتباط بين أكثر من ترتيبين (وليس ترتيبان اثنان فقط كما في حالة ارتباط سبيرمان)، كما يشيع

استخدام هذا المعامل في التحقق من ثبات الاستبيانات وغيرها حين تعرض على مجموعة من المحكمين.

- يستخدم معامل التوافق (Contingency Coefficient) في الكشف عن العلاقة أو الارتباط بين متغيرات اسمية أو نوعية (Nominal)، كما يمكن استخدامه أيضًا لقياس الارتباط بين متغير نوعي والآخر رتبي (Ordinal). ويمكن استخدام معامل التوافق سواء كان الجدول المزدوج من النمط 2×2 أو أكثر، وسواء تساوى عدد خلايا الصفوف وعدد خلايا الأعمدة أو لم يتساو.

- يستخدم كا^2 (كا تربيع) في الكثير من التطبيقات من بينها اختبار كا^2 للاستقلالية Chi Square (Test of Independence)، وفي هذه الحالة يتم التطبيق على مجموعات عينة واحدة معرفة مدى اختلافهم أو عدم اختلافهم في جانب معين (رأي أو صفة أو سلوك أو موقف أو اتجاه... إلخ)، فالهدف هنا معرفة مدى استقلالية هذا الجانب عن متغير معين ضمن خصائص العينة.إن مدى الاستقلالية يعني الاختلاف (بمعنى عدم الارتباط).

- يستخدم معامل ثيتاTheta Coefficient أو معامل فريمان (Freeman Coefficient) إذا كانت البيانات تتعلق بمتغيرين أحدهما اسمي والآخر رتبي، وعند استخدام هذا المعامل يجب أن تكون البيانات موضحة تميز تكرارات المتغير الاسمي في كل مستوى رتبي، بمعنى أننا إذا كنا بصدد بحث العلاقة بين الجنس (ذكور & إناث) ورتب مشاهدة التليفزيون،فإنه يجب تميز الذكور على الإناث أو العكس في كل رتبة.

- هناك أخطاء شائعة في تفسير معامل الارتباط، فقد يتعامل الباحث مع قيمة معامل الارتباط على أنها قيمة مطلقة مثل قيمة الطول أو الوزن مثلًا، أو على أنه

نسبة مئوية وهذا غير صحيح، كما يجب عدم تفسير معامل الارتباط على أساس وحدات الدرجات الأصلية؛لأن قيمة معامل الارتباط تكون مستقلة عن الوحدات التي يقاس بها المتغيران والقيم التي يأخذها كل منهما.

- إن دلالة معامل الارتباط هي دالة لحجم العينة، فقد يكون معامل ارتباط قدره (0.3) ولكنه دال إحصائيًا (Significant)؛ لأن حجم العينة 600 مفردة مثلًا، وفي الوقت نفسه فإن معامل ارتباط قدره (0.7) لكنه غير دال إحصائيًا (Not Significant)؛ لأن حجم العينة عشر مفردات؛ أي أن قيمة معامل الارتباط المرتفعة التي يحصل عليها الباحث باستخدام عينات صغيرة قد لا يكون لها أي دلالة إحصائية.

- إن مقدار العلاقة بين متغيرين لا تعتمد على إشارة معامل الارتباط، فمعامل الارتباط (-0.70) يعبر عن نفس مقدار العلاقة بين متغيرين معامل الارتباط بينهما (+0.70)، فالفرق بينهما يكون في اتجاه العلاقة وليس في قيمة العلاقة.

- معامل الارتباط المتعدد (R) هو الارتباط بين متغير تابع وأكثر من متغير مستقل،وتتراوح قيمة معامل الارتباط المتعدد بين الصفر والواحد الصحيح الموجب؛ أي أنه موجب القيمة دائمًا، وهذه القيمة في حقيقتها هي الجذر التربيعي لمعامل التحديد (Determination Coefficient).

- يمثل الارتباط الجزئي (Partial Correlation) أحد تطبيقات الضبط الإحصائي المهمة، وجوهر هذا الارتباط هو قياس العلاقة المستقيمة بين متغيرين بعد عزل تأثير المتغيرات الأخرى، وباستخدام معامل الارتباط الجزئي يمكن حساب معامل الارتباط بين أي متغيرين اثنين بعد تثبيت أو عزل المتغيرات الأخرى.

مصادر الفصل الثالث ومراجعه

(أ) مصادر ومراجع عربية:

- رجاء محمود أبو علام (1999)، مناهج البحث في العلوم النفسية والتربوية، (القاهرة: دار النشر للجامعات).

- زكريا الشربيني (1990)، الإحصاء اللابارامتري في العلوم النفسية والاجتماعية والتربوية، (القاهرة: مكتبة الأنجلو المصرية).

- سعدي شاكر حمودي (2000)، علم الإحصاء وتطبيقاته في المجالين التربوي والاجتماعي، (عمان: مكتبة دار الثقافة للنشر والتوزيع).

- سمير كامل عاشور& سامية أبو الفتوح سالم (2003)، العرض والتحليل الإحصائي باستخدام (SPSSWIN)، الجزء الأول، (القاهرة: جامعة القاهرة: معهد البحوث والدراسات الإحصائية).

- سمير كامل عاشور& سامية أبو الفتوح سالم (2005)، العرض والتحليل الإحصائي باستخدام (SPSSWIN)، الجزء الثاني، (القاهرة: جامعة القاهرة: معهد البحوث والدراسات الإحصائية).

- صلاح الدين محمود علام (2000)، تحليل بيانات البحوث النفسية والاجتماعية والتربوية، (القاهرة: دار الفكر العربي).

- عاطف أحمد منصور (1992)، الرياضيات المسلية، (القاهرة: مكتبة ابن سينا).

- عبد الحميد محمد نجم & محمد عبد الهادي المحميد (1990)، الإحصاء الوصفي والتحليلي مع استخدام البرامج الجاهزة، (الكويت: مكتبة جامعة الكويت).

- عزت عبد الحميد محمد حسن (2011)، الإحصاء النفسي والتربوي، (القاهرة: دار الفكر العربي).

- فتحي عبد الله فياض (1991)، التحليل الإحصائي للبيانات الجغرافية، (القاهرة: مكتبة جامعة عين شمس).

- فؤاد أبو حطب & آمال صادق (2010)، مناهج البحث وطرق التحليل الإحصائي في العلوم النفسية والاجتماعية والتربوية، (القاهرة: مكتبة الأنجلو المصرية).

- فؤاد البهي السيد (1978)، علم النفس الإحصائي، (القاهرة: دار الفكر العربي).

(ب) مصادر ومراجع أجنبية:

- Gottschalk, Louis A. (2000) The Application of Computerized Content Analysis of Natural Language in Psychotherapy Research Now and in the Future. American Journal of Psychotherapy, VOL. 54(3), pp.305-311.

- Huisman, Mark; Snijders, Tom A. B. (2003) Statistical Analysis of Longitudinal Network Data with Changing Composition. Sociological Methods & Research, Vol. 32(2), pp. 28-253.

- Janis, I. L.; Fadner, R. H.; Janowitz, M. (2005) The Reliability of A Content Analysis Technique. Public Opinion Quarterly, 7,. pp. 293-296.

- Jensun, Klaus Bruhn (2002) A Handbook of Media and Communication Research: Qualitative and Quantitative Methodologies. London. Rout Ledge.

- Kaplan, Abraham; Goldsen, Joseph M. (2005) The Reliability of Content Analysis Categories. in: Harold, D. Lasswell & Leites, Nathan, Language of Politics; Studies in Quantitative Semantics. Oxford, England: George W. Stewart, 1949. pp. 83-112. (PsycINFO Database Record (c) 2005 APA.

- Lombard, Matthew; Snyder - Duch, Jennifer; Bracken, Cheryl Campanella (2003) Content Analysis in Mass Communication: Assessment and Reporting of Inter-Coder: Reliability Correction. Human Communication Research, VOL. 29 (3), pp. 469-472.

- Mc Burney, Donald H.(2001) Research Methods. US. Wadsworth. Thomson Learning.
- Meyer, J. Patrick (2004) Background Information Methods for Statistical Analysis of Means with Mixed Groups. Dissertation Abstracts International Section A: Humanities and Social Sciences, Vol. 65 (4-A), pp. 12-52.
- Myers, Jerome L.; Well, Arnold D. (2003) Research Design and Statistical Analysis (2nd ed.). Mahwah, NJ, US: Lawrence Erlbaum Associates, Publishers.
- Rutherford, Andrew (2005) Research Design and Statistical Analysis. British Journal of Mathematical and Statistical Psychology, Vol. 58(1), pp. 187-189.
- Van Egeren, Lawrence F. (1973) Multivariate Statistical Analysis; Psychophysiology, Vol. 10 (5), pp. 517-532.

الفصل الرابع

مقاييس الفروق

تمهـــيد

يتضمن هذا الفصل تعريفًا مختصرًا ببعض الأساليب الإحصائية التي تقيس الفروق بين المجموعـات المختلفة، سواء من حيث النسب (Percentages) أو من حيث المتوسطات (Means)، ويضم الفصل ثلاثـة مباحث، يتناول المبحث الأول أساليب قياس الفروق بين مجموعتين، سواء فيما يخـص النسـب، أو فـيما يخص المتوسطات. أما المبحث الثاني، فيتناول أساليب قياس الفروق بين أكثر من مجمـوعتين مـن حيـث متوسطات الدرجات على مقياس معين، مع الأخذ في الاعتبار تعدد المتغـيرات وتفاعلهـا، في حـين يتنـاول المبحث الثالث أهم الأساليب الإحصائية اللابارامترية في قياس الفروق بين المجموعات.

المبحث الأول

الفروق بين النسب

عندما نكون بصدد معرفة معنوية الفروق بين نسـبتين (Two Percentages)، فإننـا نسـتخدم اختبـار (Z) الذي يناسب العينات الكبيرة، ولا يصلح مع العينات الصغيرة التي تقل عـن 30 مفـردة، ويسـتخدم اختبار Z لمعرفة الفروق بين النسب المستقلة، وكذلك بين النسب المرتبطة:

أولًا: دلالة الفروق بين النسب المستقلة

فيما يخص النسب المستقلة، فإنها تعني وجود مجموعتين مستقلتين (ذكور & إناث)، (ريف & حضر)، (عرب & أجانب)... إلخ، فهناك مجموعتان، لكل منهما نسبة معينة، لنفرض على سبيل المثال، أن العينة قوامها (600) مفردة، تتوزع بين الريف بواقع (245) مفردة، والحضر بواقع (355) مفردة، وأن 70.2% من مجموعة الريف يشاهدون التليفزيون بانتظام، بينما تبلغ تلك النسبة 77.7%. من الواضح اختلاف النسبتين، لكن التساؤل هو: هل توجد فروق جوهرية بين عينة الريف وعينة الحضر من حيث مشاهدة التليفزيون بانتظام؟ أم أن الفروق بينهما غير جوهرية ويمكن أن تعزى إلى الصدفة؟ للإجابة على هذا التساؤل يتم حساب معنوية الفروق بين النسب ودلالة هذه الفروق، وذلك باستخدام معادلة النسبة الحرجة، أو اختبار Z لمعرفة دلالة الفروق بين نسبتين مستقلتين، وهذه المعادلة على الصورة الآتية:

$$\text{ذ} = \frac{\text{أ}_1 - \text{أ}_2}{\sqrt{\bar{\text{أ}} \; \bar{\text{ب}} \times \frac{\text{ن}_1 + \text{ن}_2}{\text{ن}_1 \; \text{ن}_2}}}$$

حيث إن:

أ_1 نسبة الأفراد في العينة الأولى الذين أجابوا على السؤال/ البند في الوجهة المطلوبة (ولتكن نسبة الأفراد في عينة الريف الذين يشاهدون التليفزيون بانتظام).

أ_2 نسبة الأفراد في العينة الثانية الذين أجابوا على السؤال/ البند في الوجهة المعنية (ولتكن نسبة الأفراد في عينة الحضر الذين يشاهدون التليفزيون بانتظام).

$\bar{\text{أ}}$ (ألف شَرطة) هي المتوسط الموزون لنسبتي العينتين لتقدير النسبة في الأصل، ويحسب بالمعادلة:

$$\bar{أ} = \frac{ن_١ أ_١ - ن_٢ أ_٢}{ن_١ + ن_٢}$$

$\bar{ب}$ (ب شَرطة) هي ناتج طرح $\bar{أ}$ من الواحد الصحيح، أي أن:

$$\bar{ب} = 1 - \bar{أ}$$

أخيراً، فإن ($ن_1$) هي عدد مفردات المجموعة الأولى، أما ($ن_2$) فهي عدد أفراد المجموعة الثانية.

<u>مثال:</u>

في بحث عن استخدامات التليفزيون أجريت دراسة على عينة قوامها 600 مفردة، وتتوزع هذه العينة حسب منطقة الإقامة بواقع 245 من الريف مقابل 355 من الحضر، وقد كشفت الدراسة عن أن 172 من مجموعة الريف تشاهد التليفزيون بانتظام مقابل 276 من مجموعة الحضر، فهل توجد فروق جوهرية بين مجموعة الريف ومجموعة الحضر من حيث مشاهدة التليفزيون بانتظام؟

من الواضح أن نسبة مشاهدة التليفزيون بانتظام في مجموعة الريف تبلغ 70.2% أما في الحضر فتبلغ 77.7%، ويتم حساب $\bar{أ}$ بموجب المعادلة الآتية:

$$\bar{أ} = \frac{١٧٢ + ٢٧٧}{٢٤٥ + ٣٥٥} = \frac{٤٤٩}{٦٠٠} = 0.75$$

$$\bar{ب} = 1 - 0.75 = 0.25$$

ويصبح تباين النسبة المقدرة في الأصل= $0.75 \times 0.25 = 0.1875$

وباستخدام هذه المعطيات فإن:

$$ذ = \frac{٠,٧٠٢ - ٠,٧٧٧}{\sqrt{٠,١٨٧٥ \times \frac{٢٤٥ + ٣٥٥}{٢٤٥ \times ٣٥٥}}}$$

$$= \frac{-\,٠,٠٧٥}{٠,٠٣٦} = -\ 2.08$$

أي أن قيمة Z = 2.08 وهذه القيمة دالـة عنـد مسـتوى 0.05 لأنهـا تعـدت القيمـة الحرجـة 1.96 وبالتالي نقبل الفرض القائل بوجود فروق جوهرية بين عينة الريف وعينة الحضر من حيث المشاهدة المنتظمة للتليفزيون؛ حيث ترتفع نسبة المشاهدة المنتظمة في الحضر بفروق جوهرية عن الريف.

ويمكن الاستفادة من المعطيات السابقة للحصول على قيمة (Z) بموجب صيغة أخرى مبسطة للمعادلة السابقة، وهذه الصيغة هي:

$$2.08 = \frac{٠,٧٧٧ - ٠,٧٠٢}{\sqrt{٠,٧٥ \times ٠,٢٥ \left[\frac{١}{٢٤٥} + \frac{١}{٣٥٥}\right]}}$$

ونتبين أن هذه المعادلة تعتمد أيضًا على تباين النسبة المقدرة في الأصل، ويلاحظ من هذا المثال، أن عينة الريف لا تساوي عينة الحضر، $ن_1 \neq ن_2$، فبينما تبلغ عينة الريف 245 مفردة، فإن عينة الحضر تبلغ 355 غير أنه كثيرًا ما تكون المقارنة في النسبة بين عينتين متساويتين في الحجم بمعنى أن: $ن_1 = ن_2$، في مثل هذه الحالات يمكن اختصار معادلة قيمة (Z) بحيث تكون على الصورة الآتية:

$$ذ = \frac{أ_1 - أ_2}{\sqrt{\frac{٢\,\bar{أ} + \bar{ب}}{ن}}}$$

فإذا افترضنا إن عينة الريف تساوي 300 مفردة وعينة الحضر تساوي أيضًا 300 مفردة (أي أن إجمالي العينة هو 600)، وكانت نسبة مشاهدة التليفزيون بانتظام في عينة الريف هي 71.3 %؛ أي ما يعادل 214 مفردة، بينما كانت نسبة مشاهدة التليفزيون بانتظام في عينة الحضر هي 73.3 %؛ أي ما يعادل 220 مفردة، فإن:

أ ‾ = $\frac{٢١٤ + ٢٢٠}{٦٠٠}$ = $\frac{٤٤٣}{٦٠٠}$ = 0.72

ب ‾ = 1 - 0.72 = 0.28

ويصبح تباين النسبة المقدرة في الأصل= 0.72 × 0.28 = 0.2016

وباستخدام هذه المعطيات فإن:

$$ذ = \frac{٠,٧١٣ - ٠,٧٣٣}{\sqrt{\frac{٢(٠,٢٠١٦)}{٦٠٠}}} = \frac{-٠,٠٢}{٠,٠٢٦} = -0.77$$

أي أن قيمة (Z) تساوي (- 0.77) وهذه القيمة غير دالة إحصائيًا لأنها لم تصل إلى القيمة الحرجة 1.96 أي عند مستوى الثقة 0.05 وبناء على ذلك نقبل الفرض الصفري بأنه لا توجد فروق جوهرية بين عينة الريف وعينة الحضر من حيث المشاهدة المنتظمة للتليفزيون.

ثانيًا: دلالة الفروق بين النسب المرتبطة

فيما يخص النسب المرتبطة، فإنها تتعلق بمجموعة واحدة، بمعنى أن الأفراد الذين تعبر عنهم النسبة الأولى هم أنفسهم الذين تعبر عنهم النسبة الثانية، فإذا افترضنا أن لدينا عينة قوامها 400 مفردة، وأن هذه العينة تتوزع حسب مشاهدة التليفزيون بانتظام وقراءة الصحف اليومية بانتظام، فإذا كان 120 شخصًا يشاهدون التليفزيون بانتظام، كما أن 93 شخصًا يقرؤون الصحف بانتظام، فإن التساؤل هو: هل توجد فروق جوهرية بين قراءة الصحف اليومية بانتظام ومشاهدة التليفزيون بانتظام؟ أم أن الفروق بينهما غير جوهرية ويمكن أن تعزى إلى الصدفة؟ للإجابة على هذا التساؤل يتم حساب معنوية الفروق بين النسب المرتبطة ودلالة هذه الفروق، وذلك باستخدام معادلة النسبة الحرجة على الصورة الآتية:

$$ذ = \frac{أ^{١} - أ^{٢}}{\sqrt{أ^{١} + أ^{٢}}}$$

حيث أ[1] تعني عدد المفـردات التـي تعـبر عـن مشـاهدة التليفزيـون بانتظـام، أمـا أ[2] فتعنـي عـدد المفردات التي تقرأ الصحف اليومية بانتظام،وبالتعويض في المعادلة يكون:

$$ذ = \frac{١٢٠ - ٩٣}{\sqrt{١٢٠ + ٩٣}} = \frac{٢٧}{١٤,٥٩} = 1.85$$

أي أن قيمة Z هي 1.85 وهذه القيمة غـير دالـة إحصـائيّا (عنـد مسـتوى المعنويـة 0.05) لأنهـا لم تصـل إلى القيمة الحرجة وهي 1.96 ويستنتج من ذلك أنه لا توجد فروق جوهــرية بـين نسـبة قــراءة الصحف اليومية بانتظام ومشاهدة التليفزيون بانتظام.

كما يمكن استخدام $كا^{2}$ (كا تربيع) لمعرفة معنوية الفروق بين أنماط اسـتجابات عينـة واحـدة عـلى بند أو سؤال معين، لنفرض أنه في دراسة على عينة قوامها (300) مفردة، كانت آراؤهـم بشـأن الموافقـة على خصخصة التليفزيون كالآتي:

- موافق 90
- محايد 80
- رافض 130

في هذه الحالة، فإن الفرض الصفري هو (لا توجد فروق جوهرية بين هذه الأنماط الثلاثة)، بمعنـى أنه (لا توجد اختلافـات جوهريـة بـين عـدد المـوافقين والمحايـدين والرافضـين لخصخصـة التليفزيـون). وللتحقق من هذا الفرض يتم حساب التكرارات المتوقعة لكل نمط من أنماط الاستجابات، وذلك كالآتي:

التكرار المتوقع= مجمل التكرارات ÷ عدد الاستجابات

$= 300 \div 3= 100$

وتكون قيمة كا^2 = $\frac{(٩٠-١٠٠)^٢}{١٠٠} + \frac{(٨٠-١٠٠)^٢}{١٠٠} + \frac{(١٢٠-١٠٠)^٢}{١٠٠} = 9$

أي أن قيمة كا^2 المحسوبة تساوي 9 ثم نقارن هذه القيمة بقيمة كا^2 الجدولية عند درجة حرية= عدد البدائل (الاستجابات) – 1

وبما أن عدد الاستجابات هو (3)، فإن درجة الحرية = 3- 1 = 2

وبالرجوع إلى الجداول الإحصائية (جدول كا^2) نجد أنه عند درجة حرية 2، فإن قيمة كا^2 تساوي (5.99) عند مستوى الثقة 0.95 أما عند مستوى الثقة 0.99، فإن قيمة كا^2 تساوي (9.21)، ولما كانت قيمة كا^2 المحسوبة تساوي (9)، فإنها بذلك تكون أكبر من قيمة كا^2 الجدولية عند مستوى الثقة (0.95)، وبالتالي نرفض الفرض الصفري القائل بأنه (لا توجد فروق جوهرية بين هذه الأنماط الثلاثة)، وهذا الرفض يعني قبول الفرض البديل، بمعنى أن هناك فروقًا جوهرية بين عدد الموافقين والمحايدين والرافضين لخصخصة التليفزيون، ومن الواضح أن أكثرية المبحوثين عبرت عن رفض الخصخصة.

المبحث الثاني

الفروق بين متوسطين

تتم المقارنة بين متوسطين (Two Means) من خلال اختبار «ت» والمعروف اصطلاحًا (t. test) فهو من أبرز المعالجات الإحصائية لإجراء المقارنات الثنائية التي تعتمد على الوسط الحسابي. وإذا كان اختبار(t) يقيس معنوية الفروق بين متوسطين، الأول هو X^1 والثاني X^2، فإن الفرض الصفري H_0 يعني أنه لا توجد فروق جوهرية بين

هذين المتوسطين، أما الفرض البديل H_1، فيعني أنه توجد فروق جوهرية بينهما، بمعنى أنهما غير متساويين، قد يكون المتوسط الأول أقل من أو أكبر من المتوسط الثاني، وفي هذه الحالة، فإن الصيغة الإحصائية لكل من الفرض الصفري والفرض البديل هي:

$$H_0: X^1 = X^2$$

$$X^1:\ H_1 \neq X^2$$

لنفرض مثلًا أن X^1 تعني الوسط الحسابي لدرجة مجموعة الذكور على مقياس استخدام التليفزيون، أما X^2 فتعني الوسط الحسابي لدرجة مجموعة الإناث على المقياس نفسه، أما عندما تكون المقارنة بين كل من المتوسطين X^1 , X^2 لمجموعة واحدة، فإن المتوسط الأول X^1 قد يعني مثلًا متوسط درجة العينة على مقياس استخدام التليفزيون، أما المتوسط الثاني X^2 فقد يعني متوسط درجة العينة نفسها على مقياس استخدام الإنترنت. وفيما يلي توضيح أهم معادلات اختبار (t) للمقارنة بين مجموعتين أو عينتين مستقلتين، وكذلك استخدامه على مستوى العينة الواحدة.

(أ) استخدام (t. test) للمقارنة بين مجموعتين أو عينتين:

عندما تكون المقارنة بين متوسط درجة مجموعة بمتوسط درجة مجموعة أخرى، فإن هناك عدة صيغ إحصائية (معادلات) لاختبار (t) وإن كانت جميعها تقوم على المعطيات التالية:

ن1: وتعني عدد أفراد المجموعة الأولى

ن2: و تعني عدد أفراد المجموعة الثانية

ن: و تعني مجمل العينة

$\overline{\text{س}}$ 1 : وتعني متوسط درجة المجموعة الأولى

$\overline{\text{س}}$ 2: وتعني متوسط درجة المجموعة الثانية

ع^2 1: وتعني تباين درجات المجموعة الأولى

ع^2 2: وتعني تباين درجات المجموعة الثانية

ويتوقف استخدام صيغة إحصائية معينة على عدة خصائص أهمها: تساوي أو عدم تساوي عدد أفراد المجموعتين، تجانس أو عدم تجانس التباين (Variance) لدرجات المجموعتين، بمعنى اختلاف أو عدم اختلاف قيمة تباين المجموعة الأولى عن قيمة تباين المجموعة الثانية، ومع أخذ هذه العوامل بالاعتبار، فإننا نكون بصدد إحدى المقارنات الأربعة الآتية:

- إذا كانت ن1 = ن2، وتباين درجات المجموعة الأولى لا يختلف عن تباين درجات المجموعة الثانية (عدم الاختلاف هنا يعني تجانس التباين، وليس بالضرورة التساوي).
- إذا كانت ن1 = ن2، ولكن تباين درجات المجموعة الأولى يختلف عن تباين درجات المجموعة الثانية (الاختلاف هنا يعني عدم تجانس التباين).
- إذا كانت ن1 ≠ ن2، ولكن تباين درجات المجموعة الأولى = (يساوي) تباين درجات المجموعة الثانية (التساوي هنا يعني التجانس، وليس بالضرورة التساوي العام لقيمة التباين).
- إذا كانت ن1 ≠ ن2، كما أن تباين درجات المجموعة الأولى ≠ (لا يساوي) تباين درجات المجموعة الثانية (لا يساوي يعني الاختلاف، بمعنى عدم تجانس تباين العينة الأولى وتباين العينة الثانية).

وفيما يلي توضيح الصيغ الإحصائية لاختبار (t) لكل واحدة من هذه المقارنات الأربعة:

الصيغة الأولى:

إذا كانت ن1 = ن2، وتباين درجات المجموعة الأولى = (يساوي) تباين درجات المجموعة الثانية، فإن قيمة (t) تكون بموجب الصيغة الآتية:

$$\frac{\bar{س}_١ - \bar{س}_٢}{\sqrt{\frac{ع^٢_١ + ع^٢_٢}{ن - ١}}}$$

مثال: في دراسة على عينة عشوائية قوامها 500 مفردة تتوزع بالتساوي بين الذكور والإناث تبين أن متوسط درجة عينة الذكور على مقياس الاستخدامات والإشباعات هو 28 بانحراف معياري 3، بينما كان متوسط درجة عينة الإناث على المقياس نفسه 25 بانحراف معياري 4، بافتراض تجانس تباين درجات المجموعتين، هل توجد فروق جوهرية بين الذكور والإناث؟ في هذا المثال نلاحظ أن:

ن1 = 250

ن2 = 250

ع1 = 3، وبالتالي فإن $ع^2$ 1 = 9

ع2 = 4، وبالتالي فإن $ع^2$ 2 = 16

$\bar{س}$ 1 = 28

$\bar{س}$ 2 = 25

وبتطبيق المعادلة المذكورة فإن قيمة (t) =

$$\frac{٢٨ - ٢٥}{\sqrt{\frac{٩ + ١٦}{٥٠٠ - ١}}} = \frac{٣}{٠,٢٢٣٨} = 13.4$$

أي أن قيمة (t) المحسوبة تساوي 13.4 ويبقى التساؤل قائمًا عن معنوية هذه القيمة، ولحسم هذا التساؤل تتم مقارنة قيمة (t) المحسوبة بقيمة (t) الجدولية عند درجة حرية (ن – 1)، أي عند درجة حرية 499، فإذا كانت قيمة (t) المحسوبة أكبر من قيمة (t) الجدولية، كان ذلك يعني أن الفروق بين المجموعتين جوهرية ولا ترجع إلى الصدفة، أما إذا كانت قيمة (t) المحسوبة أصغر من قيمة (t) الجدولية، كان ذلك يعني أن الفروق بين المجموعتين غير جوهرية وترجع إلى الصدفة، وبالرجوع إلى الجداول الإحصائية (جدول القيم الحرجة لاختبار (t) لدلالة الطرفين عند درجة حرية 499) نجد أن قيمة t = 1.97 عند مستوى 0.05 أما عند مستوى 0.01 فإن قيمة (t) = 2.59 أي أن قيمة المحسوبة أكبر من قيمة (t) الجدولية، وعلى هذا الأساس فإن الفروق بين الجنسين جوهرية ولا ترجع إلى الصدفة.

الصيغة الثانية:

إذا كانت ن1 = ن2 مع **عدم تجانس** تباين درجات المجموعة الأولى وتباين درجات المجموعة الثانية، أي أن تباين درجات المجموعة الأولى مختلف عن تباين درجات المجموعة الثانية، في هذه الحالة تكون قيمة (t) بموجب الصيغة الآتية:

$$\frac{\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}}{\sqrt{\dfrac{\text{ن}^{١} \times \text{ع}^{٢}_{١} + \text{ن}^{٢} \times \text{ع}^{٢}_{٢}}{\text{ن}(\text{ن} - ١)}}}$$

مثال: في دراسة على عينة عشوائية قوامها 600 مفردة تتوزع بالتساوي بين الجنسين تبين أن متوسط درجة مجموعة الذكور على مقياس استخدام وسائل الإعلام هو 37 بانحراف معياري 2، بينما كان متوسط درجة مجموعة الإناث على المقياس نفسه هو 28 بانحراف معياري 5 بافتراض **عدم تجانس التباين** في المجموعتين، هل توجد فروق جوهرية بين الجنسين في استخدام وسائل الإعلام؟ في هذا المثال نلاحظ أن:

ن1 = ن2 = 300 كما أن إجمالي ن هو 600

ع1 = 2، وبالتالي فإن $ع^2$ 1 = 4 كما أن ع2 = 5 وبالتالي فإن $ع^2$ 2 = 25

$\overline{س}$ 1 = 37، بينما $\overline{س}$ 2 = 28

باستخدام هذه المعطيات، وتطبيق المعادلة السابقة، ومع الأخذ بالاعتبار **عـدم تجـانس التبـاين** في المجموعتين، فإن قيمة (t) تكون:

$$\frac{٣٧ - ٢٨}{\sqrt{\frac{٣٠٠ \times ٤ + ٣٠٠ \times ٢٥}{٦٠٠ \times ٥٩٩}}} = \frac{٩}{٠,١٥٥} = 58.06$$

أي أن قيمة (t) تساوي 58.06، وعلينا أن نقارن هذه القيمة المحسوبة بقيمة (t) الجدولية عند درجة حرية (ن – 1)، وعلى الباحث الرجوع إلى جدول القيم الحرجة لاختبار (t) لدلالة الطرفين عند درجة حرية 599 (أي 500 فأكثر)، ويتضح من ذلك الجدول أن قيمة (t) = 1.96عند مستوى 0.05 أما عند مستوى 0.01 فإن قيمة (t) = 2.58 ولما كانت قيمة (t) المحسوبة = 58.06، فإنها بذلك أكبر من قيمة (t) الجدولية، وعلى هذا الأساس فإن الفروق بين الجنسين جوهرية ولا ترجع إلى الصدفة.

الصيغة الثالثة:

إذا كانت ن1 ≠ ن2 ولكن تباين درجات المجموعة الأولى = (يساوي أو متجانس مع) تباين درجات المجموعة الثانية، فإن قيمة (t) تكون بموجب الصيغة الآتية:

$$\frac{\overline{س}_١ - \overline{س}_٢}{\sqrt{\frac{ن_١ \times ع^٢_١ + ن_٢ \times ع^٢_٢}{(ن_١ + ن_٢) - ٢}\left[\frac{١}{ن_١} + \frac{١}{ن_٢}\right]}}$$

مثال: في دراسة على عينة عشوائية قوامها 600 مفردة تتوزع بين الذكور بواقع 400 مفردة والإناث بواقع 200 مفردة، تبين أن متوسط مجموعة الذكور على مقياس استخدام وسائل الإعلام هو 58 بانحراف معياري 3، بينما كان متوسط مجموعة الإناث على المقياس نفسه هو 41 بانحراف معياري 4 بافتراض **تجانس التباين** في المجموعتين، هل توجد فروق جوهرية بين الجنسين في استخدام وسائل الإعلام؟ في هذا المثال نلاحظ أن:

ن1 = 400 بينما ن2 = 200 كما أن إجمالي ن هو 600

ع1 = 3، وبالتالي فإن $ع^2 1$ = 9 كما أن ع2 = 4، وبالتالي فإن $ع^2 2$ = 16

$\overline{س}$ 1 = 58

$\overline{س}$ 2 = 41

باستخدام هذه المعطيات، وتطبيق المعادلة السابقة، ومع الأخذ بالاعتبار **تجانس التباين** في المجموعتين، فإن قيمة (t) تكون:

$$\frac{٥٨ - ٤١}{\sqrt{\frac{٤٠٠ \times ٩ + ٢٠٠ \times ١٦}{٥٩٨}\left[\frac{١}{٤٠٠} + \frac{١}{٢٠٠}\right]}} = \frac{١٧}{٠,٢٩٢} = 58.2$$

أي أن قيمة (t) تساوي 58.2، وعلينا أن نقارن هذه القيمة المحسوبة بقيمة (t) الجدولية عند درجة حرية (ن – 2)، أي عند درجة حرية 598 (أكثر من 500)، ونتبين من جدول القيم الحرجة أن قيمة (t) عند درجة حرية أكثر من 500 تساوي = 1.96عند مستوى 0.05 أما عند مستوى 0.01، فإن قيمة t = 2.58 وبالتالي فإن قيمة (t) المحسوبة أكبر من قيمة (t) الجدولية، وعلى هذا الأساس، فإن الفروق بين الجنسين جوهرية ولا ترجع إلى الصدفة.

الصيغة الرابعة:

إذا كانت ن1 ≠ ن2 وكان تباين درجات المجموعة الأولى ≠ (لا يساوي أو غير متجانس مع) تباين درجات المجموعة الثانية، فإن حساب قيمة (t) يتم على خطوتين، الأولى حساب (ت)، والثانية حساب ($\bar{\text{ت}}$)، وتنطق ت شرطة، فيما يخص قيمة (ت)، فإن الحصول عليها يكون بموجب المعادلة:

$$\text{ت} = \frac{\bar{\text{س}}_1 - \bar{\text{س}}_2}{\sqrt{\frac{\text{ع}_1^2}{\text{ن}_1} + \frac{\text{ع}_2^2}{\text{ن}_2}}}$$

أما فيما يخص $\bar{\text{ت}}$، فإن معادلتها كالآتي:

$$\bar{\text{ت}} = \frac{\text{ت}_1 \, \text{ع}_1^2 \frac{1}{\text{ن}_1} + \text{ت}_2 \, \text{ع}_2^2 \frac{1}{\text{ن}_2}}{\sqrt{\frac{\text{ع}_1^2}{\text{ن}_1} + \frac{\text{ع}_2^2}{\text{ن}_2}}}$$

علمًا بأن: ت1 هي (ت) الجدولية بدرجات حرية (ن1 – 1) للمجموعة الأولى، بينما ت2 هي ت الجدولية بدرجات حرية (ن2 – 1) للمجموعة الثانية، وبالتالي لا بد من الحصول على (ت) الجدولية لكل مجموعة، والمقارنة بين قيمة (ت) وقيمة $\bar{\text{ت}}$:

فإذا كانت قيمة (ت) أكبر من أو تساوي $\bar{\text{ت}}$، فإن ذلك يعني أن هناك فروقًا جوهرية بين المجموعتين، أما إذا كانت قيمة (ت) أصغر من $\bar{\text{ت}}$، فإن ذلك يعني أنه لا توجد فروق جوهرية بين المجموعتين.

باستخدام هذه المعادلات المبسطة، فإن (t.test) يستخدم لاختبار معنوية الفروق بين متوسط درجات عينتين (ن1)، (ن2) (Two Sample t.test)، كالمقارنة بين متوسط

درجة الذكور ومتوسط درجة الإناث من حيث كثافة مشاهدة التليفزيون أو قراءة الصحف أو سماع الإذاعة... إلخ.

اختبار (t) لعينة واحدة:

كثيرًا ما يستخدم (t. test) على مستوى العينة الواحدة (One Sample t.test) لمعرفة معنوية الفروق بين متوسط درجة العينة على مقياس معين ومتوسط درجة العينة نفسها على مقياس آخر، أو لمقارنة متوسط درجة العينة على مقياس معين قبل وبعد إدخال مؤثر معين، أو بعد فاصل زمني بين تطبيق وآخر، وفي هذه الحالة فإن قيمة t تكون بموجب المعادلة الآتية:

$$ت = \frac{\bar{س}\ ف}{\sqrt{\frac{مجـ\ ح^{٢}ف}{ن(ن-١)}}}$$

حيث: ن: عدد مفردات العينة

$\bar{س}$ ف: متوسط الفروق بين درجات التطبيق الأول ودرجات التطبيق الثاني

مجـ $ح^2$ ف: مجموع مربعات انحرافات الفروق بين درجات التطبيقين عن الوسط الحسابي لهذه الفروق

مثال: في دراسة تجريبية على مجموعة تضم عشرة مبحوثين، كان الهدف معرفة تأثير برنامج تليفزيوني في الوعي بقضايا البيئة، وقد تم تطبيق مقياس الوعي البيئي على هؤلاء المبحوثين قبل وبعد تعرضهم للبرنامج،وتم رصد درجات المجموعة قبل وبعد تعرضهم لهذا البرنامج، وكانت على النحو التالي:

القياس القبلي: 7، 6، 7، 8، 8، 8، 9، 9، 8، 7

القياس البعدي: 10،10، 10، 10، 10، 9،10،10،9، 10

إن تطبيق اختبار (t) لمعرفة معنوية الفروق بين القياسين يتطلب حساب:

- مجموع درجات التطبيق الأول ومجموع درجات التطبيق الثاني.
- متوسط درجات التطبيق الأول ومتوسط درجات التطبيق الثاني (س ف).
- الفروق بين درجات التطبيق الأول ودرجات التطبيق الثاني (ف).
- متوسط هذه الفروق (م ف).
- انحراف الفروق عن متوسطها الحسابي، وهذا الانحراف يساوي (ف- م ف).
- مجموع مربعات الانحراف عن الفروق.

ولتوضيح هذه المتطلبات يتم تنظيم الدرجات الخام ومشتقاتها في جدول على النحو التالي:

درجات القياس القبلي	درجات القياس البعدي	ف	ح ف = ف - (-2.1)	ح ف2
7	10	3-	0.9-	0.81
6	10	4-	1.9-	3.61
7	10	3-	0.9-	0.81
8	10	2-	0.1-	0.01
8	10	2-	0.1-	0.01
8	9	1-	1.1-	1.21
9	10	1-	1.1-	1.21
9	10	1-	1.1-	1.21
8	9	1-	1.1-	1.21
7	10	3-	0.9-	0.81
مج = 77 م = 7.7	مج = 98 م = 9.8	مج ف= -21 م ف =-2.1		مج = 10.9

يتضح من هذا الجدول أن مجموع الفروق بين درجات القياس القبلي ودرجات القياس البعدي هو (-21)؛ أي أن (مج ف = - 21)، وبالتالي فإن:

(متوسط مج ف = -21 ÷ 10 = - 2.1)

هذه القيمة (-2.1) تم طرحها من فروق الدرجات، فنحصل بالتالي على (ح ف)، وهذا موضح في العمود الرابع: ح ف = ف – (-2.1)، وبتربيع قيم (ح ف)، نحصل على مربعات الانحرافات (ح ف2)، ونتبين من الجدول أن مجموعها = 10.9

ومن الجدول أيضًا فإن مجموع درجات المجموعة حسب القياس القبلي هو (77)، وبالتالي، فإن متوسط درجة القياس القبلي = 77 ÷ 10 = 7.7 أما حسب القياس البعدي فإن إجمالي درجات المجموعة هو (98) وبالتالي، فإن متوسط درجة القياس البعدي = 98 ÷ 10 = 9.8 ويكون الفرق بين المتوسطين = 9.8 – 7.7 = 2.1

باستخدام هذه المعطيات والتعويض في المعادلة المذكورة، فإن قيمة (t) تكون:

$$ت = \frac{٢,١}{\sqrt{\frac{١٠,٩}{١٠ \times ٩}}} = 6.034$$

أي أن قيمة (t) المحسوبة تساوي 6.034 وباستخدام جدول القيم الحرجة لاختبار (t) فإنه عند مقارنة هذه القيمة بقيمة (t) الجدولية عند درجة حرية (ن – 1)؛ أي عند درجة حرية (9) نجد أن قيمة (t) الجدولية = 2.26 عند مستوى 0.05 أما عند مستوى 0.01 فإن قيمة (t) الجدولية = 3.25 أي أن قيمة (t) المحسوبة أكبر من قيمة (t) الجدولية، وهذا يعني أن الفروق بين درجات العينة في القياس القبلي والقياس البعدي هي فروق جوهرية ولا ترجع إلى الصدفة، وبالتالي، فإن البرنامج التليفزيوني قد أحدث تأثيرًا جوهريًا في وعي المبحوثين بقضايا البيئة، لقد ارتفعت درجاتهم بفروق جوهرية بعد تعرضهم لهذا البرنامج.

مثال آخر: فيما يلي درجات مجموعة من المبحوثين على مقياس الإشباعات المتحققة جراء استخدام موقع (Tweeter) وموقع (Face book)، والمطلوب معرفة ما إذا كانت هناك فروق جوهرية بين الموقعين من حيث الإشباعات المتحققة:

تويتر: 12، 15، 13، 14، 14، 13، 16، 11، 11، 10

فيس بوك: 10، 13، 16، 12، 13، 12، 15، 10، 12، 9

في ضوء هذه البيانات، فإن العينة تتكون من عشرة مبحوثين، أما الانحرافات ومربعاتها، فإنها تتمثل فيما يوضحه الجدول الآتي:

إشباعات تويتر	إشباعات فيس بوك	ف	ح ف = ف - 0.7	ح ف2
12	10	2	1.3	1.69
15	13	2	1.3	1.69
13	16	3-	3.7-	13.69
14	12	2	1.3	1.69
14	13	1	0.3	0.09
13	12	1	0.3	0.09
16	15	1	0.3	0.09
11	10	1	0.3	0.09
11	12	1-	1.7-	2.89
10	9	1	0.3	0.09
مج = 129 م = 12.9	مج = 122 م = 12.2	مج ف= 7 م ف =0.7		مج = 22.1

يتضح من الجدول أن مجموع درجات الإشباعات المتحققة جراء استخدام موقع تويتر هو (129)، بمتوسط = 129 ÷ 10 = 12.9 أي أن متوسط درجة الإشباعات فيما يخص استخدام تويتر هو (12.9)، أما مجموع درجات الإشباعات المتحققة جراء استخدام

موقع فيس بوك فهو(122)، بمتوسط = 122 ÷ 10 = 12.2أي أن متوسط درجة الإشباعات فيما يخص استخدام فيس بوك هو (12.2)، ويكون الفروق بين المتوسطين = 12.9 - 12.2 = 0.7

يتضح من الجدول أيضًا أن (مج ف = 7)، وبالتالي فإن:

(متوسط مج ف = 7 ÷ 10 = 0.7)

وللحصول على (ح ف) يتم طرح هذه القيمة (0.7) من فروق الدرجات، فنحصل بالتالي على (ح ف)؛ أي انحرافات الفروق عن متوسطها الحسابي، وبتربيع قيم (ح ف)، نحصل على مربعات الانحرافات (ح $ف^2$)، ونتبين من الجدول أن مجموعها = 22.1

باستخدام هذه المعطيات والتعويض في المعادلة المذكورة، فإن قيمة (t) تكون:

$$ت = \frac{٠,٧}{\sqrt{\frac{٢٢,١}{١٠ \times ٩}}} = 1.41$$

أي أن قيمة (t) المحسوبة تساوي 1.41

وباستخدام جدول القيم الحرجة لاختبار (t)، فإنه عند مقارنة هذه القيمة بقيمة (t) الجدولية عند درجة حرية (ن – 1)؛ أي عند درجة حرية (9) نجد أن قيمة (t) الجدولية = 2.26 عند مستوى 0.05 أما عند مستوى 0.01، فإن قيمة (t) الجدولية = 3.25 أي أن قيمة (t) المحسوبة أصغر من قيمة (t) الجدولية، وهذا يعني أن الفروق بين درجات الإشباعات المتحققة جراء استخدام موقع تويتر ودرجات الإشباعات المتحققة جراء استخدام موقع فيس بوك، هي فروق غير جوهرية وترجع إلى الصدفة.

بوجه عام، فإن اختبار (t. test) يتم تطبيقه على مستوى العينة الواحدة (One Sample t.test) في الكثير من المتطلبات البحثية، كأن يكون الهدف معرفة الفروق

بين متوسطين مقترنين (Paired Sample t.test) مثال ذلك حساب معنوية الفروق بين متوسط درجة عينة على مقياس مشاهدة التليفزيون ومتوسط درجة العينة نفسها على المقياس نفسه بعد مضي فترة زمنية، أو لمعرفة معنوية الفروق بين متوسط درجة العينة على مقياس معين وبين متوسط درجة العينة نفسها على مقياس آخر، أو لمقارنة متوسط درجة العينة بمتوسط درجة المجتمع أو بمتوسط آخر مفترض، أو لمقارنة متوسط درجة العينة على مقياس معين قبل وبعد إدخال مؤثر معين.

وهذه الفكرة مهمة جدّا للبحوث التجريبية وشبه التجريبية، ففي إحدى الدراسات العلمية كان الهدف رصد وتحليل اتجاهات الطلاب نحو مهنة الخدمة الاجتماعية، وتم اختيار عينة قوامها ستون طالبًا وطالبة تتوزع بالتساوي بين الجنسين، وتم قياس اتجاهات العينة قبل تعرضها لبرنامج توجيهي يتضمن تدريبًا ميدانيًا ومشاهدة عدد من الأفلام الوثائقية التي تتناول دور الاختصاصين الاجتماعيين في المستشفيات والمدارس ومؤسسات رعاية الأيتام. وقد تلقى الطلاب هذا البرنامج، وتم قياس اتجاهاتهم نحو مهنة الخدمة الاجتماعية مرة ثانية.

في القياس الأول؛ أي قبل تلقي البرنامج - كانت شدة الاتجاه الإيجابي نحو مهنة الخدمة الاجتماعية قد جاءت بمتوسط قدرة 68.83 بانحراف معياري 21.1 أما في القياس الثاني- أي بعد تلقي البرنامج - فقد جاءت شدة الاتجاه نحو مهنة الخدمة الاجتماعية بمتوسط قدرة 84 وانحراف معياري 19.4، ويلاحظ ارتفاع قيمة المتوسط في القياس الثاني، الأمر الذي يعني أن الاتجاه الإيجابي نحو مهنة الخدمة الاجتماعية، قد ارتفع بعد البرنامج الإرشادي، لكن لا يمكن الجزم بذلك إلا بعد التأكد من أن الفروق بين المتوسطين جوهرية ولا ترجع إلى الصدفة، بمعنى أوضح، فإن التساؤل هو: هل توجد فروق جوهرية بين الاتجاه نحو مهنة الخدمة الاجتماعية قبل وبعد التدريب والتعرض للأفلام الوثائقية؟

إن تطبيق (t.test) بما يتيحه من معطيات إحصائية يمكننا من الإجابة على هذا التساؤل، علمًا بأن (ن = 60، درجة الحرية = 59)، وبتطبيق (Paired Samples t. test) جاءت قيمة (t) 6.851 وهذه القيمة دالة إحصائيّا بمستوى معنوية صفر (P = 0.000) بما يشير إلى أن الفروق جوهرية ولا ترجع إلى الصدفة، وهذا يعني أن الاتجاه الإيجابي نحو مهنة الخدمة الاجتماعية قد ارتفع ارتفاعًا جوهريّا بعد البرنامج التدريبي؛ إذ أصبح هذا المتوسط 84 بعد أن كان 68.83 قبل تعرض الطلاب للبرنامج. وكثيرًا ما يستخدم الباحثون ما يعرف بنسبة الكسب المعدل (Modified Gaining) لضمان الحصول على تقدير دقيق لفاعلية المؤثر (وهو في مثالنا المذكور التدريب المبرمج ومشاهدة الأفلام الوثائقية التي تتناول دور الاختصاصين الاجتماعيين في المستشفيات والمدارس ومؤسسات رعاية الأيتام)، ويتم الحصول على نسبة الكسب المعدل من خلال تطبيق معادلة بلاك (Blake)، وهي:

$$\frac{\text{ص} - \text{س}}{\text{د} - \text{س}} + \frac{\text{ص} - \text{س}}{\text{د}}$$

حيث:

ص: تعني درجة الاختبار البعدي

س: تعني درجة الاختبار القبلي

د: تعني النهاية العظمى لدرجة الاختبار

فإذا كانت الدرجة العظمى لمقياس الاتجاه نحو مهنة الخدمة الاجتماعية هي 100 درجة، فإن نسبة الكسب المعدل تكون:

$$\frac{\text{٨٤} - \text{٦٨,٨٣}}{\text{١٠٠} - \text{٦٨,٨٣}} + \frac{\text{٨٤} - \text{٦٨,٨٣}}{\text{١٠٠}} = 0.64$$

وإذا كان اختبار (t.test) من أكثر الاختبارات الإحصائية شيوعًا في الدراسات الاجتماعية والإنسانية، فإن هناك بعض الاحتياطات في استخدامه، وأنه بالإمكان استخدام (t.test) طالما تم التأكد من جوانب أساسية أهمها: ألا يقل عدد مفردات أي مجموعة أو عينة عن 30 مفردة، وأن يكون حجم العينتين أو المجموعتين متقاربًا (فلا يكون حجم العينة الأولى 500 مفردة بينما حجم العينة الثانية هو عشر مفردات مثلًا)؛ إذ إن التفاوت الكبير في حجم العينتين أو المجموعتين من شأنه أن يفضي إلى نتائج مضللة عند استخدام (t.test) ويرجع ذلك إلى أن دلالة قيمة (t) تتأثر بحجم العينة، (لأن درجة الحرية تعتمد على عدد مفردات العينة).

من جهة أخرى، فإن استخدام (t. test) لمعرفة معنوية الفروق بين متوسطين يتطلب إجراء اختبار تجانس التباين (Homogeneity of Variance) وتعبر عنه دلالة إحصاءات ليفن (Levine Statistics) فإذا كانت معنوية هذا الاختبار أقل من (0.05) فإن ذلك يعني عدم تجانس التباين (وهذا يفيد في تحديد صيغة معادلة اختبار (t) مع الأخذ في الاعتبار حجم العينة). ومن جهة ثالثة فإن استخدام (t.test) يتطلب التأكد أولًا من اعتدالية التوزيع في كلتا المجموعتين، ويمكن التعرف على ذلك بالعديد من الطرق منها قياس الالتواء (Skewness) لكل مجموعة، فإذا كانت الالتواء أقل من الواحد الصحيح سواء بالسالب أو بالموجب، فإن ذلك يعني أن التوزيع معتدلًا أو قريبًا من الاعتدال، وبالتالي يمكن استخدام اختبار (t.test) أما إذا كانت قيمة الالتواء تتجاوز الواحد الصحيح، فإن ذلك يتطلب تعديل التوزيع بما لا يخل بالبيانات، أو استخدام أحد بدائل اختبار (t)، ومن هذه البدائل اختبار مان وتني (Man – Whitney) وكذلك اختبار ولكوكسون (Wil Coxon Test).

المبحث الثالث

تحليل التباين أحادي الاتجاه

سبقت الإشارة إلى أننا نستخدم اختبار (t) لمعرفة معنوية الفروق بين مجموعتين فقط؛ من حيث متوسط الدرجة على مقياس معين، غير أن المقارنة قد تكون بين ثلاث مجموعات أو أكثر من حيث متوسط الدرجة. في هذه الحالة نستخدم تحليل التباين أحادي الاتجاه، (One Way Analysis of Variance) والمعروف اختصارًا (ANOVA) وهو من الأساليب الإحصائية شائعة الاستخدام في بحوث العلوم الاجتماعية والإنسانية، ويعتمد على متوسطات قيم الاستجابة على المقاييس والاختبارات المستخدمة كأدوات لجمع البيانات. ومن الجدير بالتأكيد عليه هنا أن الحديث عن التباين في هذه الجزئية يعني بنوع محدد من التباين، ألا وهو تباين الدرجات حول متوسط درجة العينة أو المجموعة، ذلك أن التباين من المنظور الإحصائي تتعدد مجالاته بصورة ملحوظة، على سبيل المثال، فإن التباين يستخدم بكثرة ضمن الانحدار، وهناك التباين المتنبأ به والتباين غير المتنبأ به (Predicted and Unpredicted Variance) فتباين الدرجات حول متوسط التوزيع هو تباين يمكن تفسيره أو التنبؤ به، أما تباين الدرجات حول خط الانحدار فهو التباين الذي لا يمكن تفسيره أو الذي لا يمكن التنبؤ به.

ويقوم تحليل التباين أحادي الاتجاه (One Way ANOVA) على المعادلات الرياضية نفسها الخاصة باختبار (t) مع الأخذ بالاعتبار تعدد المجموعات (وبالتالي تعدد المتوسطات والانحرافات المعيارية)، علمًا بأن المقارنة بين المجموعات تكون من حيث متغير تابع واحد فقط (وليس أكثر من متغير تابع)، ويمكن توضيح ذلك بمثال بسيط، لنفرض أن لدينا عينة عشوائية قوامها (500) شخص، ويتوزع هؤلاء الأشخاص حسب الحالة الاجتماعية على ثلاث مجموعات (مجموعة المتزوجين، مجموعة العُزاب، مجموعة المطلقين)، وقد تم تطبيق مقياس استخدام وسائل الإعلام على هذه العينة، وكانت متوسطات الدرجات والانحرافات المعيارية لهذه المجموعات وفق ما يوضحه الجدول الآتي:

ع	م	ن	المجموعة
2.9	34.8	103	العُزاب
6.7	31.9	347	المتزوجون
6.5	29.05	50	المطلقون
7.92	36.32	500	العينة

وكما هو واضح فإن المجموعات مستقلة عن بعضها البعض (مجموعة العُزاب، مجموعة المتزوجين، مجموعة المطلقين)، وأن لدينا متغيراً تابعًا واحـدًا (Only one Dependent Variable) ويتمثـل في الدرجـة على مقياس استخدام وسائل الإعلام، وهذا المتغـير التـابع لـه وحـدة قياسـية كميّـة (Quantitative). إن التساؤل الذي يجيب عليه تحليل التباين أحادي الاتجاه هو: هل توجد فروق جوهرية بين مجموعـات العينة في الدرجة على مقياس استخدام وسائل الإعلام حسب متغير الحالة الاجتماعية؟ بمعنى آخر: هـل يختلف متوسط درجة استخدام وسائل الإعلام من جانب المتزوجين عن العُزاب عن المطلقين؟ وإذا كـان هناك اخـتلاف جـوهري، فـأين يوجـد هـذا الاخـتلاف؟ هـل يوجـد بـين مجموعـة العُـزاب ومجموعـة المتزوجين؟ أم أنه يوجد بين مجموعة العزاب ومجموعة المطلقين؟ أم أنه يوجد بين مجموعة المتـزوجين ومجموعة المطلقين؟ أم بين هذه المجموعات الثلاث؟ إن تحليل التباين أحادي الاتجاه يجيب على هـذا التساؤل. من خلال تطبيق تحليل التباين أحادي الاتجاه يتم الحصول على معطيات إحصـائية متعـددة أهمها:

- عدد مفردات كل مجموعة (N)
- متوسط درجة كل مجموعة (Mean)
- الانحراف المعياري لمتوسط درجة كل مجموعة Standard Deviation
- الخطأ المعياري لمتوسط كل مجموعة Standard Error

هذا بالإضافة إلى درجة الثقة، والحدين الأعلى والأدنى لدرجة كل مجموعة من مجموعات العينة، متوسط درجة العينة ككل، كما تتضمن المعطيات الإحصائية نموذج تحليل التباين كاملًا ويحتوي على الاستخلاصات الأساسية ممثلة في مصدر التباين (Source of Variance)، بمعنى التباين بين المجموعات (Between Groups) والتباين داخل المجموعات (Within Groups) والتباين الكلي. فالتباين بين المجموعات يعني التباين القائم بين المجموعات الثلاثة (مجموعة المتزوجين ومجموعة العزاب ومجموعة المطلقين)، أما التباين داخل المجموعات فهو التباين بين مفردات كل مجموعة، بينما يعني التباين الإجمالي (Total) مجمل مجموع التباين داخل المجموعات ومجموع التباين بين المجموعات، كما يتضمن نموذج تحليل التباين مجموع المربعات (Sum of Squares) ومتوسط المربعات (Mean of Squares)، وكذلك درجات الحرية، وقيمة (F) ومستوى الدلالة (.Sig) والذي يعتبر حاسمًا في النموذج:

ANOVA

Source of Variance	Sum of Squares	df.	Mean of Squares	F	Sig.
Between groups	1185.111	2	592.55	9.779	000
Within groups	30116.089	497	60.596		
Total	3130102	499			

وكما واضح من هذا النموذج فإن درجة الحرية فيما يخص التباين بين المجموعات = 2 لأن هذه الدرجة تساوي (عدد المجموعات – 1)، وبقسمة مجموع المربعات بين المجموعات على 2 نحصل على متوسط المربعات (التباين) بين المجموعات. أما فيما يخص التباين داخل المجموعات، فإن درجة الحرية = 497 لأن هذه الدرجة تساوي (حجم العينة - عدد المجموعات)، وبما أن حجم العينة = 500 كما أن عدد المجموعات التي نقارن بينها = 3 فتكون درجة الحرية داخل المجموعات = 500 – 3 = 497 وبقسمة

مجموع المربعات داخل المجموعات على 497 نحصل على متوسط المربعات (التباين) داخل المجموعات، وبوجه عام، فإن قيمة متوسط المربعات (التباين) هي ناتج قسمة مجموع المربعات على درجة الحرية.

ومن النموذج أيضًا نتبين أن قيمة (F) بلغت 9.779 وهذه القيمة دالة إحصائيًا عند مستوى الصفر (Sig. = .000) وهذا يكشف بوضوح عن أن هناك فروقًا جوهرية بين مجموعات العينة في الدرجة على مقياس استخدام وسائل الإعلام (لاحظ دائمًا القيمة الموجودة تحت كلمة .Sig فإذا كانت أقل من 0.05 أو أقل من 0.01 فهذا يعني وجود فروق جوهرية، أما إذا كانت أكبر من 0.05 فهذا يعني عدم وجود فروق جوهرية)، وبالتالي يتوقف التحليل عند هذا الحد. وبما أن هناك فروقًا جوهرية، فهذا يعني استمرار التحليل لمعرفة أين توجد الفروق الجوهرية:

- هل هذه الفروق توجد بين مجموعة المتزوجين ومجموعة العُزاب؟
- أم أنها توجد بين مجموعة المتزوجين ومجموعة المطلقين؟
- أم أنها توجد بين مجموعة العُزاب ومجموعة المطلقين؟

ولا يعني ذلك أن الفروق الجوهرية تكون موجودة بين مجموعتين فقط، فقد تكون موجودة بين كل مجموعات العينة، كأن تكون هذه الفروق بين المجموعة الأولى والمجموعة الثانية، وكذلك بين المجموعة الأولى والمجموعة الثالثة، وأيضًا بين المجموعة الثانية والمجموعة الثالثة... ويتم حسم هذه المسألة من خلال أحد اختبارات المقارنات المتعددة (Multi Comparisons Test) وتتيح الحزمة الإحصائية للعلوم الاجتماعية مجموعة كبيرة من تلك الاختبارات منها:

- LSD; - Bonferroni ;- Sidak ;- Scheffe ;- R-E-G-Wf R- E – G – WQ ;- S – N – K ;- Tukey;- Tukey's – B;- Duncan ;- Hochlerg's Gt2;- Gobriel ;- Waller – Duncan - Dunnet

وبموجب تنفيذ اختبار المقارنات المتعددة، يتم الحصول على معطيات إحصائية متعددة منها مكمن الفروق الجوهرية الموجودة بين متوسطات مجموعات العينة،وفي مثالنا السابق كشف اختبار شيفيه (Scheffe) للمقارنات المتعددة عن المعطيات الموضحة بالجدول الآتي:

المقارنة	مستوى الدلالة
العُزاب -المتزوجون	.008
العُزاب - المطلقون	.000
المتزوجون -المطلقون	.086

ومن هذا الجدول نتبين بوضوح أن مستوى المعنوية بين متوسط درجة مجموعة المتزوجين ومتوسط درجة مجموعة المطلقين يبلغ 0.086، وهذه القيمة أكبر من 0.05، وبالتالي لا توجد فروق جوهرية بين المتزوجين والمطلقين من حيث الدرجة على مقياس استخدام وسائل الإعلام ($p > 0.05$) وقد أشرنا فيما سبق إلى أن متوسط درجة المتزوجين على هذا المقياس هو (31.9) درجة. أما متوسط درجة المطلقين فهو (29.05) درجة.ومن حيث المقارنة بين متوسط درجة مجموعة العُزاب ومتوسط درجة مجموعة المطلقين نتبين من الجدول أن مستوى الدلالة يساوي الصفر، أي أنه أصغر من 0.05 وبالتالي هناك فروق جوهرية بين مجموعة العُزاب ومجموعة المطلقين، وقد أشرنا فيما سبق إلى أن متوسط درجة مجموعة العُزاب هو (34.8) درجة. أما متوسط درجة مجموعة المطلقين، فهو (29.05) درجة؛ أي أن مجموعة العُزاب تستخدم وسائل الإعلام بزيادة حقيقية أو جوهرية عن مجموعة المطلقين ($p < 0.05$).

أما فيما يخص المقارنة بين متوسط درجة مجموعة العُزاب ومتوسط درجة مجموعة المتزوجين على مقياس استخدام وسائل الإعلام، فإن مستوى الدلالة يساوي 0.008؛ أي أنه أصغر كثيرًا من 0.05 وبالتالي هناك فروق جوهرية بين مجموعة العُزاب ومجموعة المتزوجين، وقد سبقت الإشارة إلى أن متوسط درجة مجموعة العُزاب هو (34.8) درجة، أما متوسط درجة مجموعة المتزوجين فهو (31.9) درجة؛ أي أن استخدام وسائل الإعلام يزداد بفروق جوهرية لدى مجموعة العُزاب عن مجموعة المتزوجين ($p < 0.05$). وعلى الباحث تفسير هذه الفروق.

معادلات تحليل التباين:

تتعدد صيغ معادلات تحليل التباين أحادي الاتجاه، وتتفاوت هذه الصيغ من حيث البساطة والتعقيد في ضوء مدى تجانس التباين بين المجموعات، وعدد هذه المجموعات، وعدد أفراد كل مجموعة. غير أنه أيّا كانت صيغة المعادلة المستخدمة في تحليل التباين، فإنها جميعًا ترتكز على حساب متوسط درجة كل مجموعة (م) ومجموع الدرجات (مج س) ومربعات هذا المجموع (مج $س^2$)، وعدد الأفراد في كل مجموعة (ن)، وكذلك مجمل أفراد هذه المجموعات، ودرجات الحرية، ومتوسطات التباين.. باستخدام هذه المعطيات يتم حساب قيمة (F) وهي النتيجة الأساسية لتحليل التباين، وذلك على النحو المبسط الذي يوضحه المثال التالي:

في دراسة تجريبية عن علاقة الشباب بالإنترنت، كانت العينة تتكون من (30) مبحوثًا يتوزعون على ثلاث مجموعات، وقد خلص الباحث إلى عدة نتائج منها ما يتعلق بالإشباعات المتحققة جراء استخدام موقع التواصل الاجتماعي تويتر (Tweeter) والجدول الآتي يوضح درجات أفراد كل مجموعة من حيث الإشباع المتحقق:

المجموعة الأولى	المجموعة الثانية	المجموعة الثالثة
8	9	7
10	7	8
10	6	10
10	5	8
10	7	7
8	7	6
9	6	6
9	6	8
8	-	7
7	-	6
-	-	5
-	-	8

فإذا أراد الباحث معرفة ما إذا كانت هناك فروق جوهرية بين المجموعات الثلاث، من حيث متوسط الدرجة على مقياس الإشباعات أم لا، فإن عليه أولًا الحصول على المعطيات الأساسية التالية:

(1) تحديد درجات الحرية: وذلك في ضوء حجم العينة، وعدد المجموعات التي تتم المقارنة بينها (وفي مثالنا المذكور فإن العينة تضم ثلاثين مفردة موزعة على ثلاث مجموعات، بواقع عشرة مبحوثين في المجموعة الأولى، ثمانية مبحوثين في المجموعة الثانية، اثنا عشر مبحوثًا في المجموعة الثالثة)، ويتم تحديد درجة الحرية كالآتي:

- درجة الحرية بين المجموعات تساوي: (عدد المجموعات – 1) وفي مثالنا المذكور لدينا ثلاث مجموعات، وبالتالي فإن درجة الحرية بين المجموعات تساوي: 3- 1 = 2
- درجة الحرية داخل المجموعات تساوي: (حجم العينة – عدد المجموعات)، وفي مثالنا المذكور لدينا ثلاثين مفردة موزعة على ثلاث مجموعات، وبالتالي فإن درجة الحرية داخل المجموعات تساوي: 30- 3 = 27

(2) حساب متوسط درجة كل مجموعة، وذلك بقسمة إجمالي درجات كل مجموعة على عدد أفرادها، ومن الجدول السابق، فإن إجمالي درجات المجموعة الأولى هو (89)، وعدد أفرادها (10)، فيكون متوسط درجة هذه المجموعة = 89 ÷10 = 8.9 أما متوسط درجة المجموعة الثانية فهو (6.6) بينما متوسط درجة المجموعة الثالثة (7.2)

(3) حساب مجموع الدرجات (مج س): من الجدول السابق فإن مجموع درجات المجموعة الأولى (89)، أما مجموع درجات المجموعة الثانية، فهو (53) درجة،

بينما مجموع درجات المجموعة الثالثة (86) درجة، ويكون مجموع درجات كل أفراد العينـة هو (228) أي أن (مج س) = 228

(4) حساب مجموع مربعات الدرجات (مج س2): وذلك من خلال مربع درجة كل فرد، ثم جمـع هذه المربعات للحصول مجموع مربعات درجات العينة كلها. إن ذلك يكون بضرب درجة كـل مبحوث في نفسها، وجمع الناتج على النحو المبين بالجدول الآتي:

مربعات درجات أفراد المجموعة الأولى	مربعات درجات أفراد المجموعة الثانية	مربعات درجات أفراد المجموعة الثالثة
64	81	49
100	49	64
100	36	100
100	25	64
100	49	49
64	49	36
81	36	36
81	36	64
64	-	49
49	-	36
-	-	25
-	-	64
مج س2= 803	مج س2= 361	مج س2= 636

ويتضح من هذا الجدول أن مجموع مربعات درجات المجموعات الثلاث هو (803+ 361+ 636) = 1800 أي أن (مج س2) يساوي (1800)

بموجب ذلك يتم تحليل التباين وفق الخطوات الأساسية الآتية:

(1) حساب (مجموع المربعات الكلي): ويكون ذلك بتطبيق المعادلة الآتية:

$$\text{مجموع المربعات الكلي} = \text{مج س}^2 - \frac{(\text{مج س})^٢}{\text{ن}}$$

$$= 1800 - \frac{(٢٢٨)^٢}{٣٠} = 67.2$$

(2) حساب مجموع المربعات بين المجموعات، ومعادلته:

$$\frac{(\text{مج س}١)^٢}{\text{ن}١} + \frac{(\text{مج س}٢)^٢}{\text{ن}٢} + \frac{(\text{مج س}٣)^٢}{\text{ن}٣} - \frac{(\text{مج س})^٢}{\text{ن}}$$

حيث إن: مج س تعني مجمل درجة المجموعة، كما أن مج س^2 مربع مجمل درجة كل مجموعة، (مج س)^2 تعني مربع مجموع درجات العينة.

بالتعويض في المعادلة المذكورة، فإن التباين بين المجموعات =

$$\frac{(٨٩)^٢}{١٠} + \frac{(٥٣)^٢}{٨} + \frac{(٨٦)^٢}{١٢} - \frac{(٢٢٨)^٢}{٣٠}$$

= 792.1 + 351.125 + 616.3 – 1732.8 = 26.725

أي أن مجموع المربعات بين المجموعات = 26.725

ويكون التباين بين المجموعات = مجموع المربعات بين المجموعات

÷ (عدد المجموعات -1)

= 26.725 ÷ 2 = 13.4

أي أن التباين بين المجموعات يساوي (13.4)

(لاحظ أن «عدد المجموعات- 1» هو درجة الحرية بين المجموعات)

(3) حساب مجموع المربعات داخل المجموعات، وهو يساوي:

مجموع المربعات الكلي – مجموع المربعات بين المجموعات

40.4 = 26.725 – 67.2 =

أي أن مجموع المربعات داخل المجموعات = 40.4 ومن ثمّ يتم حساب التباين داخل المجموعات، هذا التباين يساوي:

مجموع المربعات داخل المجموعات ÷ (حجم العينة- عدد المجموعات)

1.5 = 27 ÷ 40.4 =

أي أن التباين داخل المجموعات = 1.5

(لاحظ أن «حجم العينة - عدد المجموعات» هو درجة الحرية داخل المجموعات)

(3) حساب قيمة (F) وذلك بموجب المعادلة:

التباين بين المجموعات ÷ التباين داخل المجموعات = 13.4 ÷ 1.5 = 8.9

أي أن قيمة (F) المحسوبة تساوي (8.9).

(4) مقارنة قيمة (F) المحسوبة بقيمة (F) الجدولية عند درجة حرية 2 للبسط، ودرجة حرية 27 للمقام. وبالرجوع إلى الجداول الإحصائية نجد أن قيمة (F) الجدولية = 3.35 عند مستوى 0.05 أما عند مستوى 0.01 فإن قيمة (F) الجدولية = 5.49 أي أن قيمة (F) المحسوبة أكبر من قيمة (F) الجدولية عند مستوى (0.01)؛ أي أن هناك فروقًا جوهرية بين مجموعات العينة من حيث متوسط الدرجة على المقياس المستخدم.

ويتم تنظيم معطيات الخطوات السابقة في جدول مختصر على النحو الآتي:

مصدر التباين	درجة الحرية	مجموع المربعات	متوسط المربعات (التباين)	قيمة F
بين المجموعات	2	26.758	13.4	
داخل المجموعات	27	40.4	1.5	**8.9
المجموع	29	46.7		

** دالة عند مستوى 0.01

ونظرًا لوجود فروق جوهرية بين المجموعات من حيث متوسط الدرجة على المقياس، فإن على الباحث أن يستخدم إحدى طرق المقارنات البعدية، وعلى الرغم من تعدد طرق المقارنات البعدية كما سبقت الإشارة، إلا أن الأكثر استخدامًا هي: اختبار شيفيه (Scheffe)، وطريقة أقل الفروق (LSD)، وطريقة توكي (Tukey)؛ حيث تستخدم هذه الطرق كثيرًا في بحوث الاتصال والبحوث النفسية والاجتماعية والتربوية. وفي مثالنا السابق كشف تحليل التباين عن وجود فروق جوهرية بين المجموعات الثلاث، وفيما يلي توضيح كيفية التعرف على مكمن هذه الفروق باستخدام أبرز طرق المقارنات البعدية:

(أ) اختبار شيفيه للمقارنات المتعددة (Scheffe Multi-Comparison Test):

يتم تطبيق هذا الاختبار وفق الخطوات المبسطة التالية:

(1) حساب الوسط التوافقي: وذلك لأن المجموعات التي نقارن بينها تختلف من حيث عدد الأفراد، أما إذا كان عدد الأفراد متساويًا في المجموعات الثلاث، فلا يتم استخدام الوسط التوافقي. ويرتكز هذا الوسط على عدد المجموعات التي نقارن بينها وعدد أفراد كل مجموعة، ويتم الحصول على الوسط التوافقي بموجب المعادلة:

$$\frac{\text{عدد المجموعات}}{\frac{١}{ن_{١}} + \frac{١}{ن_{٢}} + \frac{١}{ن_{٣}} + \frac{١}{ن_{..X}}}$$

حيث: ن1 هي المجموعة الأولى، ن2 هي المجموعة الثانية، ن3 هي المجموعة الثالثة، نx هـي آخـر مجموعة، بموجب ذلك، فإنه في مثالنا لدينا ثـلاث مجموعـات، كـل مجموعـة تضـم عـددًا محـددًا مـن الأفراد، ويكون الوسط التوافقي يساوي:

$$9.7 = \frac{٣}{\frac{١}{١٢} + \frac{١}{٨} + \frac{١}{١٠}}$$

(تم حساب الوسط التوافقي بهذه الطريقـة؛ لأن أعـداد أفـراد المجموعـات غـير متسـاوية، أمـا إذا كانت أعداد أفراد المجموعات متساوية، كأن يكون عدد أفـراد العينـة ثلاثـين شخصًـا تتـوزع بالتسـاوي بواقع عشرة أفراد لكـل مجموعـة- فإنـه يـتم اسـتخدام عـدد أفـراد أي مجموعـة كبـديل عـن الوسـط التوافقي).

(2) تطبيق معادلة اختبار (Scheffe) مع الاستعانة بقيمة (F) الجدولية عند درجة حرية التباين بين المجموعات، ودرجة حرية التباين داخل المجموعات؛ أي عند درجة حرية 2 ودرجـة حريـة 27 وكما سبقت الإشارة، فإن قيمة (F) الجدولية = 3.35 عند مستوى 0.05 أما عند مسـتوى 0.01 فإن قيمة (F) الجدولية = 5.49

فإذا استخدمنا مستوى الثقة (0.95)، فإن معادلة اختبار (Scheffe) للمقارنات المتعددة تكون:

$$SCH = \sqrt{\frac{\text{2 (التباين داخل المجموعات) (عدد المجموعات – 1) (قيمة F الجدولية)}}{\text{الوسط التوافقي}}}$$

$$= \sqrt{\frac{٢(١,٥)(٢)(٣,٣٥)}{٩,٧}} = 1.44$$

ومن الواضح أنه تم التعويض عـن (F) الجدوليـة بالقيمـة (3.35) وهـي قيمـة (F) الجدوليـة عنـد مستوى الثقة (0.95)، في حين تم التعويض عن التباين داخل المجموعات بالقيمة 1.5

أما إذا استخدمنا مستوى المعنوية 0.01 فإن قيمة (F) الجدولية = 5.49 وعلى ذلك تكون قيمة

$$\text{اختبار (Scheffe)} = \sqrt{\frac{٢(١,٥)(٢)(٥,٤٩)}{٩,٧}} = 1.84$$

وقد تم التعويض عن (F) بالقيمة (5.49) وهي قيمة (F) الجدولية عند مستوى الثقة (0.99)، كما تم التعويض عن التباين داخل المجموعات بالقيمة 1.5

(3) مقارنة الفروق بين المتوسطات بقيمة اختبار (Scheffe) ويكون الفرق بين أي متوسطتين غير جوهري إذا كانت قيمته تقل عن (1.44)، بينما يكون هذا الفرق جوهريّا عند مستوى 0.05 إذا تراوحت قيمته ما بين 1.44 لأقل من 1.84 في حين يكون الفرق بين أي متوسطتين جوهريّا عند مستوى 0.01 إذا بلغت قيمة هذا الفرق 1.84 أو أكثر، وبناء على ذلك يتعين حساب الفروق بين متوسطي درجات كل مجموعتين نقارن بينهما على النحو التالي:

الفرق بين م ن 1، م ن 2 (8.9 – 6.6 = 2.3)

الفرق بين م ن 1، م ن 3 (8.9 – 7.2 = 1.7)

الفرق بين م ن 2، م ن 3 (7.2 – 6.6 = 0.6)

(4) وأخيرًا، يتم تنظيم هذه الفروق في جدول ثنائي/تبادلي على النحو التالي:

المجموعة م =	**الأولى م = 8.9**	**الثانية م = 6.6**	**الثالثة م = 7.2**
الأولى م = 8.9	-		
الثانية م = 6.6	2.3**	-	
الثالثة م = 7.2	1.7*	0.6	-

**$p < 0.01$

*$p < 0.01$

ونتبين من هذا الجدول بسهولة أنه:

- توجد فروق جوهرية بين المجموعة الأولى والمجموعة الثانية، بدرجة ثقة 0.99 والفرق لصالح المجموعة الأولى (م = 8.9) مقارنة بالمجموعة الثانية (م = 6.6)
- توجد فروق جوهرية بين المجموعة الأولى والمجموعة الثالثة بدرجة ثقة 0.95 والفرق لصالح المجموعة الأولى (م =8.9) مقارنة بالمجموعة الثالثة (م = 7.2)
- لا توجد فروق جوهرية بين المجموعة الثانية والمجموعة الثالثة؛ إذ إن الفروق بين متوسطي المجموعتين أقل من قيمة اختبار (Scheffe).

(ب) طريقة أقل الفروق دلالة (Least Significant Differences): وتعرف اختصارًا (LSD)، وتعتمد هذه الطريقة على:

- الوسط التوافقي (وقد سبق حسابه، وقيمته 9.7)
- متوسط التباين داخل المجموعات (وقد سبق حسابه، وقيمته 1.5)
- درجة الحرية داخل المجموعات، وهي 27 كما سبقت الإشارة
- القيمة الجدولية لقيمة (t) لدلالة الطرفين عند درجة حرية (27)، وبالرجوع إلى جدول القيم الحرجة لاختبار (t) عند درجة حرية (27)، فإن قيمة (t) عند مستوى 0.05 = 2.05 أما عند مستوى 0.01 فإنها تساوي 2.77

وعلى ذلك يتم حساب أقل الفروق الدالة (LSD) عند مستوى (0.05) بموجب المعادلة:

$$2.05 \sqrt{\text{متوسط التبادين داخل المجموعات} \left[\frac{2}{\text{الوسط التوافقي}} \right]}$$

$$2.05\sqrt{١,٥\left[\frac{٢}{٩,٧}\right]} = 1.14$$

أما قيمة أقل الفروق الدالة (LSD) عند مستوى (0.01) فهو يكون بموجب المعادلة:

$$2.77\sqrt{١,٥\left[\frac{٢}{٩,٧}\right]} = 1.54$$

(3) مقارنة الفروق بين المتوسطات بقيمة اختبار (LSD) ويكون الفرق بين أي متوسطتين غير جوهري إذا كانت قيمته تقل عن (1.14)، بينما يكون هذا الفرق جوهريًا عند مستوى 0.05 إذا تراوحت قيمته ما بين 1.14 لأقل من 1.54، في حين يكون الفرق بين أي متوسطتين جوهريًا عند مستوى 0.01 إذا بلغت قيمة هذا الفرق 1.54 أو أكثر

وبالنظر إلى الفروق بين متوسطات درجات المجموعات على النحو السابق ذكره، فإنه تتأكد النتيجة نفسها والتي تم التوصل إليها باختبار (Scheffe)؛ حيث توجد فروق جوهرية بين المجموعة الأولى والمجموعة الثانية، بدرجة ثقة 0.99 وذلك لصالح المجموعة الأولى، كما توجد فروق جوهرية بين المجموعة الأولى والمجموعة الثالثة بدرجة ثقة 0.95 وذلك لصالح المجموعة الأولى، وأخيرًا فإنه لا توجد فروق جوهرية بين المجموعة الثانية والمجموعة الثالثة؛ إذ إن الفروق بين متوسطي المجموعتين يساوي (0.6)، وهو أقل من قيمة اختبار أقل الفروق دلالة.

(ج) طريقة توكي (Tukey):

تعتمد هذه الطريقة على المعطيات اللازمة للاختبارين السابقين، بالإضافة إلى قيمة (Q) الجدولية (والتي يتم الحصول عليها من جدول القيم الحرجة لاختبار توكي)، وبالرجوع إلى هذا الجدول، نجد أن قيمة (Q) عند درجة حرية (2) ودرجة حرية (27) عند مستوى 0.05 = 2.77 أما عند مستوى 0.01 فإنها تساوي 3.64

وعلى ذلك يتم حساب قيمة اختبار (Tukey) بموجب المعادلة الآتية:

$$Q \text{ الجدولية } \sqrt{\frac{\text{التباين داخل المجموعات}}{\text{الوسط التوافقي}}}$$

فإذا استخدمنا مستوى المعنوية (0.05) فإن قيمة Tukey تكون:

$$1.09 = \sqrt{\frac{١,٥}{٩,٧}} \; 2.77$$

أما إذا استخدمنا مستوى المعنوية (0.01) فإن قيمة (Tukey) تكون:

$$1.43 = \sqrt{\frac{١,٥}{٩,٧}} \; 3.64$$

(3) مقارنة الفروق بين المتوسطات بقيمة اختبار (Tukey) يكون الفرق بين أي متوسطتين غير جوهري إذا كانت قيمته تقل عن (1.08)، بينما يكون هذا الفرق جوهريًا عند مستوى 0.05 إذا تراوحت قيمته ما بين 1.08 لأقل من 1.43 في حين يكون الفرق بين أي متوسطتين جوهريّا عند مستوى 0.01 إذا بلغت قيمة هذا الفرق 1.43 أو أكثر

وبالنظر إلى الفروق بين متوسطات درجات المجموعات على النحو السابق ذكره، فإنه تتأكد النتيجة السابقة؛ حيث توجد فروق جوهرية بين المجموعة الأولى والمجموعة الثانية، بدرجة ثقة 0.99 لصالح المجموعة الأولى، كما توجد فروق جوهرية بين المجموعة الأولى والمجموعة الثالثة بدرجة ثقة 0.95 لصالح المجموعة الأولى، ولا توجد فروق جوهرية بين المجموعة الثانية والمجموعة الثالثة.

اعتبارات أساسية في تحليل التباين:

هناك مجموعة من الاعتبارات الأساسية عند استخدام تحليل التباين، ومع التسليم بتعدد هذه الاعتبارات، إلا أن أهمها يتعلق بــ:متى يتوقف تحليل التباين؟ واستخدام

المتوسط الوزني، واستخدام الوسط التوافقي (Harmonic Mean)، وفيما يلي توضيح هذه النقاط بإيجاز.

(أ) متى يتوقف تحليل التباين؟:

يقصد بتوقف تحليل التباين عدم إجراء المقارنات المتعددة، بمعنى عدم استخدام أيٍّ من أساليب الاختبارات البعدية. إن عدم استخدام تلك الأساليب يكون إذا كانت قيمة (F) المحسوبة أقل من قيمة (F) الجدولية، ففي هذه الحالة لا تكون هناك فروق بين المجموعات، ويتوقف تحليل التباين عند هذا الحد.

مثال: في دراسة تجريبية عن علاقة الشباب بالإنترنت، كانت العينة تتكون من (30) مبحوثًا يتوزعون على ثلاث مجموعات، وقد خلص الباحث إلى عدة نتائج، منها ما يتعلق بالإشباعات المتحققة جراء استخدام موقع التواصل الاجتماعي تويتر (Tweeter)، والجدول الآتي يوضح درجات أفراد كل مجموعة من حيث الإشباع المتحقق:

المجموعة الأولى	المجموعة الثانية	المجموعة الثالثة
8	9	10
7	7	8
7	6	7
6	5	6
5	7	6
8	7	8
9	6	7
9	6	6
8	7	5
7	8	8

ففي هذا المثال.. وباتباع الخطوات السابق ذكرها، فإن:

- مجموع الدرجات (مج س)، وبحساب ذلك من الجدول، فإن مجموع درجات المجموعة الأولى هو (74)، أما مجموع درجات المجموعة الثانية، فهو (68)، في حين مجموع درجات المجموعة الثالثة (71)؛ أي أن (مج س= 213).

- مجموع مربعات الدرجات: بتربيع درجات كل مبحوث، وجمع ناتج ذلك على مستوى كل مجموعة، وبتنفيذ هذه الخطوة، فإن مجموع مربعات درجات المجموعة الأولى هو (562)، ومجموع مربعات درجات المجموعة الثانية هو (474)، أما مجموع مربعات درجات المجموعة الثالثة فهو (523)، وبالتالي فإن (مج س 2) = 562+ 474+ 523 = 1559 أي أن (مج س2) يساوي (1559)

- متوسط درجة المجموعة الأولى = 7.4 أما متوسط درجة المجموعة الثانية فهو (6.8) بينما متوسط درجة المجموعة الثالثة (7.1)، مع ملاحظة أن كل مجموعة تضم عشرة مبحوثين.

- درجة الحرية بين المجموعات = 3- 1 = 2

- درجة الحرية داخل المجموعات = 30- 3 = 27

بموجب ذلك يتم تحليل التباين وفق الخطوات السابق ذكرها على النحو التالي:

(1) حساب التباين الكلي (مجموع المربعات الكلي): ويكون ذلك بتطبيق المعادلة الآتية:

$$1559 - \frac{(٢١٣)^{٢}}{٣٠} = 46.7$$

(2) مجموع المربعات بين المجموعات =

$$\frac{(٧٤)^{٢}}{١٠} + \frac{(٦٨)^{٢}}{١٠} + \frac{(٧١)^{٢}}{١٠} - \frac{(٢١٣)^{٢}}{٣٠}$$

$$= 547.6 + 462.4 + 504.1 = 1514.1 - 1512.3 = 1.8$$

وبالتالي يكون التباين بين المجموعات = 1.8 ÷ (3-1) = 0.9

(3) مجموع المربعات داخل المجموعات، وهو يساوي:

= 46.7 – 1.8 = 44.9

وبالتالي يكون التباين داخل المجموعات = 44.9 ÷ (30- 3) = 1.7

(4) قيمة (F) = 0.9 ÷ 1.7 = 0.529

أي أن قيمة (F) المحسوبة = 0.529 وعلينا أن نقارن قيمة (F) المحسوبة بقيمة (F) الجدولية، وبالرجوع إلى الجداول الإحصائية (جدول الدلالة الإحصائية لقيمة (F)) نتبين أن قيمة (F) عند درجة حرية 27 للتباين الكبير، ودرجة حرية 2 للتباين الصغير = 3.35 وذلك عند مستوى الثقة 0.95أي أن قيمة (F) المحسوبة أصغر من قيمة (F) الجدولية، وبالتالي لا توجد فروق جوهرية بين المجموعات من حيث متوسط الدرجة على المقياس المستخدم، ويتوقف تحليل التباين عند هذا الحد.

(ب) المتوسط الوزني:

الجزئية الأخرى الجديرة بالتوضيح هي استخدام المتوسط الوزني (Weighted Mean) في حساب مجموع المربعات بين المجموعات (وبالتالي التباين بين المجموعات)، لقد سبق شرح كيفية حساب هذا المجموع، غير أن هناك طريقة أخرى تتمثل في استخدام المتوسط الوزني لحساب مجموع المربعات بين المجموعات، ألا وهي طريقة المتوسط الوزني (م و)، ومعادلته:

$$\text{م و} = \frac{\text{ن}^{١}\,\text{م}^{١} + \text{ن}^{٢}\,\text{م}^{٢} + \text{ن}^{٣}\text{م}^{٣}}{\text{ن}^{١} + \text{ن}^{٢} + \text{ن}^{٣}}$$

حيث ن 1 م1 عدد أفراد المجموعة الأولى مضروبًا في متوسط درجاتها

ن 2 م2 عدد أفراد المجموعة الثانية مضروبًا في متوسط درجاتها

ن 3 م3 عدد أفراد المجموعة الثالثة مضروبًا في متوسط درجاتها

فإذا افترضنا أن المجموعات التي نقارن بينها هي ثلاث مجموعات، كل مجموعة عشرة أفراد، وكان متوسط درجة المجموعة الأولى (7.4)، ومتوسط درجة المجموعة الثانية (6.8)، ومتوسط درجة المجموعة الثالثة (7.1)، فإن المتوسط الوزني:

$$\text{م و} = \frac{(١٠\times٧,٤) + (١٠\times٦,٨) + (١٠\times٧,١)}{١٠+١٠+١٠} = 7.1$$

ونظرًا لأن أعداد الأفراد في المجموعات متساوية (حيث: ن1 = ن2 = ن3)، يمكن اختصار هذه المعادلة كالآتي:

م و= 10 (7.4 + 6.8+ 7.1) ÷ 30 = 7.1

أي أن المتوسط الوزني هو (7.1)

وباستخدام المتوسط الوزني (م و) يكون مجموع المربعات بين المجموعات =

ن1 (م1- م و)^2 + ن2 (م2- م و)^2 + ن3 (م3- م و)^2

= 10(7.4- 7.1)^2 + 10(6.8- 7.1)^2 + 10(7.1- 7.1)^2

= 10× (0.09) + 10(0.09) + 10(صفر)

= 0.9 + 0.9 + صفر = 1.8

ويكون التباين بين المجموعات = 1.8 ÷ 2 = 0.9

وهي النتيجة نفسها التي تم التوصل إليها عندما تم حساب التباين بالطريقة العادية (السابق توضيحها)، دون استخدام المتوسط الوزني.

(ج) الوسط التوافقي:

سبقت الإشارة إلى أننا نستخدم الوسط التوافقي في اختبارات المقارنات المتعددة (Multi Comparison Tests) ويقوم الوسط التوافقي على عدد المجموعات التي نقارن بينها وعدد أفراد كل مجموعة، وإذا كانت المجموعات التي نقارن بينها غير متساوية من حيث عدد أفراد كل مجموعة، فإن الحصول على الوسط التوافقي يكون بموجب المعادلة السابق ذكرها وهي:

$$\frac{\text{عدد المجموعات}}{\frac{١}{ن_١} + \frac{١}{ن_٢} + \frac{١}{ن_٣} + \frac{١}{ن..x}}$$

حيث: ن1 هي المجموعة الأولى، ن2 هي المجموعة الثانية، ن3 هي المجموعة الثالثة، نx هي آخر مجموعة، وأن كل مجموعة تختلف من حيث عدد الأفراد (ن1 ≠ ن2 ≠ ن3)، فقد تكون المقارنة بين ثلاث مجموعات، الأولى تضم (50) مفردة، والثانية تضم (55) مفردة، في حين تضم الثالثة (47) مفردة... فكل مجموعة تختلف عن الأخرى من حيث عدد المفردات، في هذه الحالة يتعين الحصول على الوسط التوافقي بموجب المعادلة المذكورة.

غير أن المقارنة قد تكون بين مجموعات متساوية؛ حيث (ن1 = ن2 = ن3)، كأن تضم المجموعة الأولى (50) مفردة، والثانية تضم (50) مفردة، والثالثة تضم (50) مفردة، في هذه الحالة لا نستخدم الوسط التوافقي، ويتم التعويض عنه بقيمة (ن)، بمعنى عدد أفراد أي مجموعة طالما أن (ن1 = ن 2 = ن3)، وبالتالي يتم حساب اختبار شيفيه بموجب المعادلة:

تكون:

$$SCH = \sqrt{\frac{2\ \text{(التباين داخل المجموعات) (عدد المجموعات} - 1\text{) (قيمة F الجدولية)}}{\text{ن}}}$$

التفاعل بين المتغيرات المستقلة:

في الجزئية السابقة كانت المقارنة بين مجموعات العينة مصنفة حسب متغير مستقل واحد، ألا وهو متغير الحالة الاجتماعية، لكن متوسط الدرجة على المقياس المستخدم قد يختلف حسب متغيرين معًا (وليس حسب كل متغير على حدة)، كما أن هذا المتوسط قد يختلف حسب ثلاثة متغيرات مستقلة أو أكثر. فيما يخص التفاعل بين متغيرين مستقلين، مثال ذلك: أن يختلف متوسط الدرجة على مقياس استخدام التليفزيون حسب متغير الجنس ومتغير الحالة الاجتماعية (معًا)، وهنا نستخدم التفاعل الثنائي (Bilateral Interaction)، وإذا صنفنا عينة الدراسة حسب هذين المتغيرين، فإن عدد مجموعات العينة يتضاعف، فحسب متغير الحالة الاجتماعية بمفرده كان لدينا ثلاث مجموعات (أعزب، متزوج، مطلق)، أما إذا أخذنا في الاعتبار متغير الجنس (ذكور وإناث)، فسوف يكون لدينا ست مجموعات هي:

- مجموعة الذكور العُزاب
- مجموعة الإناث العزباوات
- مجموعة الذكور المتزوجين
- مجموعة الإناث المتزوجات
- مجموعة الذكور المطلقين
- مجموعة الإناث المطلقات

أي أن كل مجموعة حسب الحالة الاجتماعية تم تقسيمها إلى مجموعتين حسب الجنس، أو أن كل مجموعة حسب الجنس تم تقسيمها إلى ثلاث مجموعات حسب الحالة الاجتماعية:

الحالة الاجتماعية	الجنس	
	ذكور	إناث
أعزب		
متزوج		
مطلق		

لنفرض أن التساؤل هو: هل يختلف متوسط درجة مجموعة الذكور المتزوجين عن متوسط درجة مجموعة الإناث المتزوجات على مقياس مشاهدة التليفزيون؟ يلاحظ من هذا التساؤل أننا أخذنا في الاعتبار متغيرين مستقلين معا، هما متغير الجنس ومتغير الحالة الاجتماعية، ويعبر عن ذلك إحصائيًا بالتفاعل الثنائي بين المتغيرات المستقلة من حيث تأثير هذا التفاعل على متغير تابع؛ أي أن التعدد هنا إنما هو على مستوى المتغيرات المستقلة، أما المتغير التابع فهو متغير واحد فقط (Only One Dependent Variable)، وهذا يختلف عن تحليل التباين متعدد الاتجاهاتMANOVA حيث يكون هناك أكثر من متغير تابع (More Than One Dependent Variable).

والواقع أن استخدام أكثر من متغير في تصنيف مجموعات العينة يكون عادة بعد المقارنة بين المجموعات مصنفة حسب متغير واحد؛ من حيث متوسط الدرجة على مقياس معين (أيّا كان عدد هذه المجموعات)، ثم يريد الباحث أن يتعمق بعض الشيء لمعرفة أثر التفاعل بين المتغيرات المستقلة على المتغير التابع.

كمثال توضيحي لهذه الفكرة نفرض أن إحدى الدراسات تقصت التوجه نحو المهنة لدى عينة من القائمين بالاتصال (ن= 928)، وعند المقارنة بين المجموعات الفرعية

باستخدام اختبار (t) وتحليل التباين (ANOVA) اتضح أن استجابات المفحوصين على مقياس التوجه نحو المهنة جاءت بمتوسط قدره 10.1 وانحراف معياري 3.7، وتختلف قيمة متوسط التوجه نحو المهنة حسب المتغيرات الديموجرافية والمهنية على النحو المبين بالجدول الآتي:

المتغيرات	ن	م	ع	مؤشرات إحصائية
الجنس: ذكور	282	9	3.6	T=6.1*
إناث	646	10.6	3.7	
التعليم: جامعي	833	10.2	3.8	T=1.4
فوق الجامعي	95	9.6	3.6	
مدة الخبرة: أقل من 10 سنوات	524	10.1	3.6	T=0.06
10 سنوات فأكثر	404	10.2	3.8	
الدخل: منخفض	387	10.1	3.8	F =1.1
متوسط	280	10.4	3.6	
مرتفع	261	9.9	3.8	
العينة	928	10.1	3.7	

$^*P < 0.05$

واضح من هذا الجدول عدم وجود فروق بين مجموعات العينة من حيث التوجه نحو المهنة وذلك حسب متغيرات: مدة الخبرة ومتغير المستوى التعليمي والدخل (P>0.05)، لكن هناك فروقًا جوهرية في التوجه نحو المهنة لدى مجموعات العينة حسب متغير الجنس (P<0.05)؛ حيث يتضح من الجدول أن الإناث أكثر توجهًا نحو المهنة مقارنة بالذكور (م=10.6 للإناث مقابل 9 للذكور)، وإذا كان ذلك فيما يخص معنوية الفروق بين مجموعات العينة في التوجه نحو المهنة حسب المتغيرات الموضحة بالجدول كل متغير على حدة، فإنه باستخدام التفاعل الثنائي(Bilateral Interaction) بين هذه المتغيرات خلصت الدراسة إلى النتيجة الموضحة بالجدول الآتي:

مصدر التباين	درجة الحرية	مجموع المربعات	متوسط المربعات	حجم التأثير	قيمة F
الجنس× التعليم	1	11.4	11.4	0.001	0.85
الجنس×مدة الخبرة	1	3.6	3.6	صفر	0.3
الجنس× الدخل	2	61.5	30.7	0.005	2.3
التعليم ×مدة الخبرة	1	88.7	88.7	0.007	*6.4
التعليم × الدخل	2	16	8	0.001	0.6
مدة الخبرة×ا لدخل	2	62.6	31.3	0.005	2.4

*$P < 0.05$

يتضمن هذا الجدول ستة تفاعلات ثنائية، هي كافة التفاعلات الثنائية الممكنـة بـين المتغـيرات (وذلـك على أساس أن لدينا أربعة متغيرات مستقلة، وبالتالي يكون عدد التفاعلات الثنائية الممكنـة هـو 4× 3 ÷2= 6 ومن بين هذه التفاعلات الستة يوجد تفاعل واحد فقط ذو دلالة إحصائية للدرجة على مقيـاس التوجـه نحو المهنة، ألا وهو التفاعل بين متغير المستوى التعليمي ومتغير مدة الخبرة (وهذه نتيجة جديـدة هامـة تم الحصول عليها) باسـتخدام التفاعـل الثنـائي؛ ونظـرًا لوجـود دلالـة لهـذا التفاعـل، فإنـه يتعـين تقصـي متوسطات درجات مجموعات العينة حسب هذين المتغـيرين معًا (متغير المستوى التعليمي ومتغير مـدة الخبرة)، وباستخدام التفاعل ثنائي الاتجـاه جاءت النتيجـة على النحو المبين بالجدول الآتي:

التعليم	مدة الخبرة			
	أقل من 10 سنوات		10 سنوات فأكثر	
	م	ع	م	ع
جامعي	10.2	3.7	10.1	3.8
فوق الجامعي	8.6	3.4	10.5	3.6

يتضمن الجدول أربع مجموعات فرعية تضمها العينة حسب متغير مدة الخبرة ومتغير التعليم معًا:

- مجموعة المفحوصين ذوي التعليم الجامعي والخبرة الأقل من 10 سنوات
- مجموعة المفحوصين ذوي التعليم فوق الجامعي والخبرة الأقل من 10 سنوات
- مجموعة المفحوصين ذوي التعليم الجامعي والخبرة 10 سنوات فأكثر
- مجموعة المفحوصين ذوي التعليم فوق الجامعي والخبرة 10 سنوات فأكثر

وقرين كل مجموعة متوسط الدرجة (م) والانحراف المعياري (ع)، ومن الجدول نتبين بوضوح أن القائمين بالاتصال ذوي الخبرة الأطول في العمل، الحاصلين على مؤهلات فوق الجامعية هم أكثر مجموعات العينة تعبيرًا عن التوجه الإيجابي نحو المهنة (م=10.5)، يليهم مجموعة القائمين بالاتصال ذوي الخبرة الأقل من عشر سنوات الحاصلين على مؤهل جامعي (م= 10.2) وفي المرتبة نفسها تقريبًا تأتي مجموعة القائمين بالاتصال ذوي الخبرة الأكثر من عشر سنوات الحاصلين على مؤهل جامعي (م= 10.1)، أما أقل المجموعات تعبيرًا عن هذا التوجه – حسب مدة الخبرة والمستوى التعليمي – فهي مجموعة القائمين بالاتصال ذوي الخبر الأقل من عشر سنوات الحاصلين على مؤهلات فوق الجامعية (م=8.6).. هكذا يتضح لنا أن استخدام التفاعل ثنائي الاتجاه قد مكننا من رصد متوسطات درجات مجموعات العينة على مقياس التوجه نحو المهنة حسب مدة الخبرة والمستوى التعليمي معًا.

وقد يكون التفاعل بين ثلاثة متغيرات مستقلة أو أكثر، وهو يعرف بالتفاعل المتعدد (Multi Interaction)، كمثال على ذلك عندما تستهدف الدراسة رصد متوسطات درجات مجموعات القائمين بالاتصال على مقياس التوجه نحو المهنة حسب متغيرات الجنس والمستوى التعليمي ومدة الخبرة (معًا)، وحسب هذه المتغيرات الثلاثة يكون لدينا المجموعات الآتية:

الجنس		ذكور		إناث	
التعليم		جامعي	فوق الجامعي	جامعي	فوق الجامعي
الخبرة	أقل من 10				
	10 سنوات فأكثر				

- مجموعة الذكور ذوي التعليم الجامعي، ذوي الخبرة الأقل من 10 سنوات
- مجموعة الذكور ذوي التعليم الجامعي، ذوي الخبرة 10 سنوات فأكثر
- مجموعة الذكور ذوي التعليم فوق الجامعي، ذوي الخبرة الأقل من 10 سنوات
- مجموعة الذكور ذوي التعليم فوق الجامعي، ذوي الخبرة 10سنوات فأكثر
- مجموعة الإناث ذوات التعليم الجامعي، ذوات الخبرة الأقل من 10 سنوات
- مجموعة الإناث ذوات التعليم الجامعي، ذوات الخبرة 10 سنوات فأكثر
- مجموعة الإناث ذوات التعليم فوق الجامعي، ذوات الخبرة الأقل من 10 سنوات
- مجموعة الإناث ذوات التعليم فوق الجامعي، ذوات الخبرة 10 سنوات فأكثر

أي أنه حسب متغيرات الجنس والتعليم ومدة الخبرة، فإن العينة تضم ثمان مجموعات فرعية، ومن الطبيعي أن يزداد عدد مجموعات العينة كلما زاد عدد المتغيرات التي ستؤخذ في الاعتبار في التفاعل ونحن بصدد تطبيق تحليل التباين أحادي الاتجاه، كما أن زيادة عدد المتغيرات التي يتضمنها التفاعل، يتطلب أن تكون العينة كبيرة بما يضمن وجود العدد الكافي من المفردات في كل مجموعة فرعية، وفي الوقت نفسه يضمن عدم الإخلال بشرط أساسي في استخدام المقارنات، ألا وهو التقارب بين عدد مفردات المجموعات التي تتم المقارنة بينها، فلا يعقل مثلًا أن تكون إحدى المجموعات 100 مفردة ومجموعة أخرى خمس مفردات.

المبحث الرابع

أساليب لابارامترية لقياس الفروق

سبقت الإشارة إلى معنى الإحصاء اللابارامتري، والذي يتضمن بدائل متعددة لقياس الفروق بين المجموعات، فاختبار مان ويتني (U Test) (Mann-Whitney) مثلًا هو بديل لابارامتري لاختبار (t)، كما أن اختبار كروسكال- واليس (Kruskal- Wallis Test) هو بديل لابارامتري لتحليل التباين ذي الاتجاه الواحد (One-Way ANOVA).. وهكذا. غير أنه قبل توضيح أهم الأساليب الإحصائية اللابارامترية الخاصة بقياس الفروق، يتعين توضيح المقصود بمصطلح «المجموعات المستقلة» مقابل مصطلح «المجموعات غير المستقلة»، من الضروري التنويه إلى المعنى الدقيق لكلا المصطلحين؛ لأن هناك معاملات إحصائية تستخدم مع مصطلح معين دون الآخر؛ ومن ثمّ فإن الخلط بين المصطلحين يترتب عليه استخدامات خاطئة للأساليب الإحصائية.

إن المقصود بـ«المجموعات المستقلة» هي المجموعات المتمايزة في الصفة أو الخاصية المقاسة، فالذكور مجموعة متمايزة، والإناث مجموعة متمايزة؛ أي أننا بصدد مجموعتين مستقلتين، وعندما نصنف عينة من المبحوثين حسب محل الإقامة إلى (ريف، حضر، بدو)، فإننا نكون أمام مجموعات مستقلة أو متمايزة.

أما «المجموعات غير المستقلة»، فيقصد بها المجموعات المرتبطة، بمعنى أن المجموعة الأولى هي نفسها المجموعة الثانية، أو متطابقة معها، ويتضح ذلك على سبيل المثال في تطبيق المقياس على مجموعة من المبحوثين ثم إعادة تطبيقه مرة أخرى على المجموعة نفسها، ففي هذه الحالة نكون بصدد مجموعتين مرتبطتين (غير مستقلتين)، كما أن المجموعات المرتبطة (غير المستقلة) تكون أيضًا في حالة أن تكون مفردات العينة عبارة عن أزواج متناظرة أو متطابقة (Matched Pairs). وهنا يتعين الكشف عن اتجاه وحجم الفروق بينها.

وتستخدم الأساليب الإحصائية اللابارامترية في التحقـق مـن الفـروض الفارقـة ذات المتغـيرات الاسـمية والمتغيرات الرتبية، وتختلف هذه الأساليب حسب عدد المجموعات (العينات) وطبيعتها (مسـتقلة أو غـير مستقلة)، وكذلك حسب طبيعة المتغيرات (اسمية أو رتبية) على النحو المبين الذي يلخصه هذا الجدول:

المجموعات (العينات)	البيانات	
	اسمية	رتبية
عينـــة واحدة	- اختبار العشوائية - اختبار ذو الحدين - كا^2 - اختبار كولموجروف سمير نوف	
مجموعـتان مستقلتان	- اختبار فيشر - كا^2 لمقارنة مجموعتين - اختبار الوسيط	- اختبار كولموجروف سمير نوف - اختبار مان وتني
مجموعـتان غير مستقلتين	- اختبار ماكنمارا	- اختبار ولكوكسون - اختبار الإشارة
ثلاث مجموعات (أو أكثر) مستقلة	- كا^2 للاستقلالية - اختبار الوسيط	- كروسكال واليس
ثلاث مجموعات (أو أكثر) غير مستقلة	- اختبار كوكران	- اختبار فريدمان

وفـيما يـلي توضـيح للأسـاليب الإحصـائية اللابارامتريـة الموضـحة بالجـدول مـن خلال تصنيفها تحت تقسيمين أساسيين: الأول أساليب إحصائية لقيـاس الفـروق في حالـة البيانات الاسمية، أما الثاني فهو أساليب إحصائية لقياس الفروق في حالة البيانات الرتبيـة،

مع الأخذ في الاعتبار عدد المجموعات وطبيعتها؛ من حيث ما إذا كانت مجموعات مستقلة أو مجموعات غير مستقلة:

أولًا: قياس الفروق في حالة البيانات الاسمية

(أ) بيانات اسمية لعينة واحدة:

(1) اختبار عشوائية الدورة (Run Test): يقوم هذا الاختبار على مفهوم الدورة (Run) بمعنى دورة الحدوث، فالدورة هي مجموعة الأحداث المتشابهة، وقد يسبقها أو يعقبها أحداث مخالفة لها، كما قد لا يسبقها ولا يعقبها أحداث (Events) مخالفة. إن عدد الأحداث في دورة الحدوث يطلق عليه طول الدورة. ويختص (Run Test) بمعرفة ما إذا كان حدوث أو ظهور قيمة معينة لمتغير ثنائي يتم عشوائيًا أم لا؛ أي أن المتغير ثنائي التصنيف (ذكور & إناث)... فإذا رمزنا لهذه الثنائية بالرقم (1) والرقم (2)، فإن المتغير قد يأخذ الرقم (1) أو الرقم (2)، التساؤل هو ما إذا كانت القيمة التي يأخذها المتغير تحدث عشوائيًا أم لا؟ لنفترض أن لدينا مجتمعًا يتساوى فيه عدد الإناث مع عدد الذكور، وأننا سحبنا عينة من هذا المجتمع لإجراء دراسة تستهدف معرفة معدل استخدام الإنترنت، هنا يمكن استخدام اختبار عشوائية الدورة (Run Test) لمعرفة ما إذا كانت هذه العينة عشوائية أم غير عشوائية، فإذا كان معظم مفردات العينة من جنس واحد (ذكور فقط أو إناث فقط)، فإن عشوائية العينة تكون محل تساؤل أو شكوك.

(2) اختبار ذو الحدين (Binomial Test): يستخدم هذا الاختبار لمعرفة معنوية الفروق بين نسبتين لمتغيرات اسمية ثنائية التقسيم (نعم- لا)، (يشاهد- لا يشاهد)، (يقرأ الصحف- لا يقرأ الصحف)، فالباحث قد يريد معرفة ما إذا كانت هناك فروق جوهرية بين عدد المشاهدين وعدد غير المشاهدين، أو بين عدد

القراء وعدد غير القراء في إطار العينة الواحدة التي تتوزع مفرداتها حسب المشاهدة وعدم المشاهدة، أو بين القراءة وعدم القراءة... إلخ.

(3) اختبار كا^2 (x^2) (Chi Square Test): يستخدم هذا الاختبار للمقارنة بين التكرارات المشاهدة (Observed) والتكرارات المتوقعة (Expected)، ويستخدم مع البيانات الاسمية ثلاثية التقسيم ضمن العينة الواحدة(كأن تكون الوسيلة المفضلة: الصحف اليومية، التليفزيون، الإنترنت)، ويريد الباحث التحقق من الفرض القائل (لا توجد فروق جوهرية بين الوسائل الإعلامية من حيث التفضيل).

(4) اختبار كولموجروف - سمير نوف (Kolmogrov-Smirnov Test)، ويستخدم نفس استخدام كا^2 المشار إليه خاصة إذا كان حجم العينة أكبر من 30 (ن > 30)، ففي هذه الحالة يكون اختبار كولموجروف – سمير نوف أكثر دقة من اختبار كا^2.

(ب) بيانات اسمية لمجموعتين مستقلتين:

(1) اختبار فيشر (Fisher Test): يستخدم هذا الاختبار للمقارنة بين مجموعتين مستقلتين (ذكور & إناث أو غير ذلك)؛ من حيث متغير اسمي ثنائي (مثال ذلك: مشاهدة نشطة & مشاهدة غير نشطة)، وذلك لاختبار الفرض الصفري القائل بعدم وجود فروق بين الجنسين في نمط المشاهدة (كمثال).

(2) اختبار كا^2 (x^2) (Chi Square Test) لمقارنة مجموعتين:

يستخدم هذا الاختبار أيضًا للمقارنة بين مجموعتين مستقلتين؛ من حيث متغير اسمي (Nominal) كالمقارنة بين مجموعة الذكور ومجموعة الإناث؛ من حيث

نمط قضاء الوقت (مشاهدة التليفزيون- قراءة الصحف- قراءة الكتب)، أو للمقارنة بين مجموعة المصريين ومجموعة الكويتيين؛ من حيث الوسيلة الإعلامية المفضلة (تليفزيون الدولة- القنوات الأجنبية).

(3) اختبار الوسيط (The Median Test): يستخدم هذا الاختبار للمقارنة بين مجموعتين من حيث الوسيط (Median) في متغير اسمي معين (كالمقارنة بين الذكور والإناث من حيث وجود صفة معينة لها صفة المتغير الاسمي (الاهتمام بالسياسة & عدم الاهتمام بالسياسة)، ففي هذه الحالة يكون هناك تكرارات للذكور وأخرى للإناث، ويتم إيجاد الوسيط للمجموعتين، ثم نحدد عدد الأفراد الذين تقع تكراراتهم فوق الوسيط، وعدد الأفراد الذين تقع تكراراتهم تحت الوسيط أو مساوية له لكلٍّ من الذكور والإناث؛ ومن ثمّ نعرف معنوية الفروق بين الجنسين.

(ج) بيانات اسمية لمجموعتين غير مستقلتين:

وهنا يستخدم اختبار ماكنمارا (McNamara Test) الذي يقيس معنوية الفروق بين تطبيقين من حيث متغير اسمي ثنائي التقسيم، مثال ذلك المقارنة بين عدد المجموعة التي تشاهد برنامج تليفزيوني معين في الفترة الصباحية وعدد المجموعة نفسها التي تشاهد هذا البرنامج في الفترة المسائية، أو المقارنة بين قراءة الصحيفة اليومية من جانب عينة البحث في يوم معين بقراءتها في يوم آخر من جانب العينة نفسها.

(د) بيانات اسمية لثلاث مجموعات مستقلة أو أكثر:

(1) اختبار كا^2 للاستقلالية (Chi-Square Test for Independence): وهو يستخدم للتحقق من عدم وجود فروق جوهرية بين ثلاث مجموعات مستقلة (أو أكثر)، من حيث متغير اسمي (كالفروق بين مجموعات: الريف، الحضر، البدو)؛ من حيث نوعية البرامج المفضلة (الأخبار، الدراما، برامج المناقشات، البرامج الدينية)، فالتعامل هنا يكون مع التكرارات الخاصة بكل مجموعة.

(2) اختبار الوسيط (The Median Test): إذا كـان اختبـار الوسـيط يسـتخدم للمقارنـة بـين عينتـين مستقلتين كما سبقت الإشارة، فإنه يستخدم أيضًا للمقارنة بين ثلاث عينـات مسـتقلة (أو أكـثر من ثلاث عينات)، فمن خلال اختبار الوسيط يمكن المقارنة بين ثـلاث مجموعـات مسـتقلة (أو أكثر)؛ من حيث الوسيط (Median) في متغير اسمي معين، كالمقارنة بين مجموعـات: (القـاهرة، الوجه القبلي، الوجه البحـري)؛ مـن حيـث الصـحف المقـروءة (الأهـرام، الأخبـار، الجمهوريـة، الوفد، المصري اليوم)، وذلك اعتمادًا على التكرارات موزعة على هذين المتغيرين.

(هـ) بيانات اسمية لثلاث مجموعات (أو أكثر) غير مستقلة:

يستخدم في ذلك اختبار كوكران (The Cochran Q Test) ويتطلب تطبيـق هـذا الاختبـار أن تكـون الاستجابات ثنائية ويتم تكويدها بالصفر أو الواحـد الصـحيح، فـإذا كانـت الاسـتجابات هـي مشـاهدة إعلان معين، فإن المشاهدة تأخذ الرقم الكودي (1) أما عدم المشـاهدة فتأخـذ الـرقم الكـودي (صـفر). مثال: لدينا مجموعة من المبحوثين وقد أجريت عليهم دراسة تستهدف معرفـة مـدى مشـاهدة إعـلان تليفزيوني معين ثلاث مرات، وفي كل مرة كان الإعلان يذاع بشكل مختلف، ونريـد التحقـق مـن الفـرض التالي: «لا توجد فروق دالة إحصائيًا في مشاهدة الإعلان التليفزيوني باختلاف الشكل الذي يعـرض بـه». إن اختبار كوكران يوضح لنا ما إذا كانت هناك فروق جوهريـة في مشـاهدة الإعـلان باختـلاف الشـكل الذي يعرض به في التليفزيون.

ثانيًا: قياس الفروق في حالة البيانات الرتبية

(أ) الفروق الرتبية لمجموعتين مستقلتين:

(1) اختبار كولموجروف - سمير نـوف (Kolmogrov-Smirnov Test) وهـو يسـتخدم لاختبـار دلالـة الفـروق بـين مجمـوعتين مسـتقلتين؛ مـن حيـث متغـير رتبـي

(كالمقارنة بين الذكور والإناث من حيث تفضيل الإنترنت: لا أفضل- أفضل بدرجة ضعيفة- أفضل بدرجة متوسطة- أفضل بدرجة مرتفعة).

(2) اختبار مان- وتني (Mann-Whitney) (U Test): يستخدم هذا الاختبار للمقارنة بين عينتين مستقلتين (ذكور & إناث) أو (ريف & حضر).... وذلك في متغير رتبي، أو في متغير آخر بعد تحويله إلى متغير رتبي، وهذا الاختبار يستخدم كبديل عن اختبار (T) عندما لا تتوافر شروط اختبار (T) أو عندما يصعب تطبيق اختبار (T) لسبب أو لآخر.

(ب) الفروق الرتبية لمجموعتين غير مستقلتين:

يتيح الإحصاء اللابارامتري أساليب إحصائية متنوعة لمعرفة ما إذا كانت هناك فروق جوهرية بين متوسطين لمتغير رتبي معين، وذلك لمجموعتين غير مستقلتين،ومن هذه الأساليب:

(1) اختبار ولكوكسون (Wilcoxon Test): ويستخدم هذا الاختبار لمعرفة مقدار الفروق بين مجموعتين مرتبطتين (غير مستقلتين)؛ وذلك من حيث متغير رتبي (كالمقارنة بين مستوى درجات مجموعة من المبحوثين على مقياس كثافة مشاهدة التليفزيون (منخفض- متوسط- مرتفع) قبل أحداث معينة بدرجاتهم على المقياس نفسه بعد وقوع هذه الأحداث، أو مقارنة مستوى تفضيل عينة من الناخبين لمرشح معين قبل تنفيذ الحملة الدعائية وبعد تنفيذ هذه الحملة، أو لمقارنة متوسط رتب درجة عينة الثبات على مقياس معين في التطبيق الأول بمتوسط درجة العينة نفسها على المقياس نفسه في التطبيق الثاني، أو لمقارنة متوسط رتب درجة العينة على مقياس استخدام التليفزيون بمتوسط رتب درجتها على مقياس استخدام الصحف اليومية... إلخ.

(2) اختبـــار الإشـــارة (Sign Test): إن الإشـــارة هنـــا يقصـــد بهـــا الزيـــادة (+) أو النقصـــان (-)؛ أي أن هذا الاختبار يستخدم لمعرفـة الإشـارة (اتجـاه الفـروق بـين عينتـين غـير مسـتقلتين بالزيادة أو بالنقصان) في متغير رتبي معين، كالمقارنة بين مستوى تأييد مرشح معين من جانـب عينة من الناخبين منذ شهر مـثلًا ومسـتوى تأييـدهم لـه الآن (هـل ارتفـع مسـتوى التأييـد أم انخفض؟)، ولا يبين لنا اختبار الإشارة مقدار التغير زيـادة أو نقصًـا، ولكنـه يبـين اتجـاه التغـير فقط (ويمكن إجراء اختبار الإشارة مع البيانات الفترية بالإضافة إلى البيانات الرتبية).

(ج) الفروق الرتبية لثلاث مجموعات (أو أكثر) مستقلة:

في هذه الحالة يستخدم اختبار كروسـكال- والـيس (Kruskal- Wallis Test) وهـو بـديل لابـارامتري لتحليل التباين ذي الاتجاه الواحـد (One-Way ANOVA)؛ حيـث يسـتخدم اختبـار كروسـكال- والـيس لمعرفة الفروق بين المجموعات المستقلة (ثلاث مجموعات أو أكـثر) مـن حيـث متغـير رتبـي معـين، ولا يشترط تساوي عدد مفردات المجموعات. ويقوم هذا الاختبار على إعطاء رتب لدرجة كل فرد من أفراد المجموعات التي تتم المقارنة بينها، ومن هذه الرتب يتم معرفة الفروق بين هذه المجموعات، وإذا كان تحليل التباين (ANOVA) يستخدم لمعرفة معنوية الفروق بين ثـلاث مجموعـات (أو أكـثر) مـن حيـث متوسط الدرجة على مقياس فتري (Interval)، فإن اختبار كروسكال- والـيس يسـتخدم لمعرفـة معنويـة الفروق بين ثلاث مجموعات (أو أكثر) من حيث متوسط الدرجة على مقياس رتبي.

(د) الفروق الرتبية لثلاث مجموعات (أو أكثر) غير مستقلة:

عندما نريد الكشف عن معنوية الفروق بين بيانات رتبية لـثلاث مجمــوعات غـير مســتقلة (أو أكــثر) يسـتخدم اختبـار فريـدمان (The Friedman Test)؛ حيـث يمكـن

معرفة مدى وجود فروق جوهرية بين ترتيب ثلاثة بدائل (أو أكثر من ثلاثة بـدائل) وذلـك حسـب آراء العينة. تصور أن الفرض يقول: لا توجد فروق جوهرية بين درجة تفضيل وسائل الإعلام (صحافة، راديو، تليفزيون). إن اختبار فريدمان يكشف لنا عما إذا كانت هناك فـروق دالـة إحصـائيّا بـين ترتيـب هـذه الوسائل الثلاث؛ من حيث الأفضلية حسب آراء العينة. وفي الدراسات التجريبية، فإن العينة قد تتعـرض لثلاثة مواقف مختلفة، وتحصل على تقييم عقب كل موقف، فيكون للمبحوث الواحـد ثلاثـة تقيـيمات موضحة بالرتب (Ranks)، وهنا يمكن استخدام اختبار فريدمان لمعرفة مدى وجود فروق جوهريـة بـين الدرجات الرتبية التي حصل عليها المبحوثون في المواقف الثلاثة.

هذه بإيجاز شديد أهم الأساليب الإحصائية اللابارامترية لقياس الفروق، ويمكن الحصول على المزيد من المعلومات عن هذه الأساليب وبرامجها الإحصائية مـن الكتيبـات الإلكترونيـة (Electronic Manuals) التي تتضمنها البرامج الإحصائية الحاسوبية، كما يمكن الرجوع إلى مجلة الإحصـاء اللابـارامتري (Journal of Nonparametric Statistics) التي تصدرها الرابطة الإحصائية الأمريكية وموقعها الإلكتروني هو:

www.amstat.org/section/bus_econ/zellner.htmL

خلاصة الفصل الرابع

تناول هذا الفصل تعريفًا مختصرًا ببعض الأساليب الإحصائية التي تقيس الفروق بين مجموعات العينة ويتلخص في:

- النسب المستقلة تعني وجود مجموعتين مستقلتين، بمعنى أن هناك مجموعتين، لكل منهما نسبة معينة،أما النسب المرتبطة (غير المستقلة) فتتعلق بمجموعة واحدة، بمعنى أن الأفراد الذين تعبر عنهم النسبة الأولى هم أنفسهم الذين تعبر عنهم النسبة الثانية (كأن يكون 80% من العينة يشاهدون التليفزيون، مقابل 93% من العينة نفسها يستخدمون الفيسبوك)، ولقياس معنوية الفروق بين النسب (سواء المستقلة أو المرتبطة) نستخدم معادلة النسبة الحرجة، أو اختبار (Z)، كما يمكن استخدام كا^2 (كا تربيع) وذلك بصفة خاصة لمعرفة معنوية الفروق بين أنماط استجابات عينة واحدة على بند أو سؤال معين.

- تتم المقارنة بين متوسطين (Two Means) من خلال اختبار «ت» والمعروف اصطلاحًا (t. test)، وهو من أهم المعاملات الإحصائية لإجراء المقارنات الثنائية التي تعتمد على الوسط الحسابي. والفرض الصفري H_0 يعني أنه لا توجد فروق جوهرية بين هذين المتوسطين، أما الفرض البديل H_1، فيعني أنه توجد فروق جوهرية بينهما، ويستخدم اختبار (t) للمقارنة بين مجموعتين أو عينتين مستقلتين، وكذلك استخدامه على مستوى العينة الواحدة.

- -هناك عدة صيغ إحصائية (معادلات) لاختبار (t) ويتوقف استخدام صيغة إحصائية معينة على عدة خصائص أهمها: تساوي أو عدم تساوي عدد أفراد المجموعتين، تجانس أو عدم تجانس التباين (Variance) لدرجات المجموعتين،

بمعنى اختلاف أو عدم اختلاف قيمة تباين المجموعة الأولى عن قيمة تباين المجموعة الثانية.

- يقتضي استخدام (t. test) التأكد من جوانب أساسية أهمها ألا يقل عدد مفردات أي مجموعة أو عينة عن 30 مفردة، وأن يكون حجم العينتين أو المجموعتين متقاربًا (فلا يكون حجم العينة الأولى 500 مفردة، بينما حجم العينة الثانية عشر مفردات مثلًا)؛ إذ إن التفاوت الكبير في حجم العينتين أو المجموعتين من شأنه أن يفضي إلى نتائج مضللة عند استخدام (t.test)، ويرجع ذلك إلى أن دلالة قيمة (T) تتأثر بحجم العينة (لأن درجة الحرية تعتمد على عدد مفردات العينة).

- إن استخدام (t. test) لمعرفة معنوية الفروق بين متوسطين يتطلب إجراء اختبار تجانس التباين (Homogeneity of Variance) والتأكد من اعتدالية التوزيع في كلتا المجموعتين، بمعنى أن يكون التوزيع معتدلًا أو قريبًا من الاعتدال.

- يستخدم تحليل التباين أحادي الاتجاه (One Way Analysis of Variance) والمعروف اختصارًا (ANOVA) للمقارنة قد تكون بين ثلاث مجموعات أو أكثر؛ من حيث متوسط الدرجة على مقياس معين أو متغير تابع، علمًا بأن المقارنة بين المجموعات تكون من حيث متغير تابع واحد فقط (وليس أكثر من متغير تابع).

- تتعدد صيغ معادلات تحليل التباين أحادي الاتجاه، وتتفاوت هذه الصيغ من حيث البساطة والتعقيد في ضوء مدى تجانس التباين بين المجموعات، وعدد هذه المجموعات، وعدد أفراد كل مجموعة. وأيّا كانت صيغة المعادلة المستخدمة في تحليل التباين، فإنها جميعًا ترتكز على حساب متوسط درجة كل مجموعة (م) ومجموع الدرجات (مج س) ومربعات هذا المجموع (مج س2)، وعدد الأفراد في كل

مجموعة (ن)، وكذلك مجمل أفراد هذه المجموعات، ودرجات الحرية، ومتوسطات التباين.

- يتم تحليل التباين وفق خطوات أساسية هي: حساب (مجموع المربعات الكلى)، حساب مجموع المربعات بين المجموعات، حساب مجموع المربعات داخل المجموعات، ومن ثمّ حساب قيمة = (F) (وهي تساوي التباين بين المجموعات ÷ التباين داخل المجموعات)، ومقارنة قيمة (F) المحسوبة بقيمة (F) الجدولية، وتكون هناك فروق جوهرية بين المجموعات إذا كانت قيمة (F) المحسوبة أكبر من قيمة (F) الجدولية، بينما لا تكون هناك فروق جوهرية بين المجموعات إذا كانت قيمة (F) المحسوبة أقل من أو تساوي قيمة (F) الجدولية.

- إذا تبين عدم فروق جوهرية بين المجموعات، فإن تحليل التباين يتوقف عند هذا الحد، أما إذا تبين وجود فروق جوهرية بين المجموعات، فإنه يتم استخدام إحدى طرق المقارنات البعدية، مثل اختبار شيفيه (Scheffe)، وطريقة أقل الفروق (LSD)، وطريقة توكي (Tukey) لمعرفة أي المجموعات توجد بينها فروق جوهرية.

- التفاعل بين المتغيرات المستقلة يعني تأثير أكثر من متغير مستقل في المتغير التابع، مثال ذلك أن يختلف متوسط الدرجة على مقياس استخدام التليفزيون حسب متغير النوع ومتغير الحالة الاجتماعية (معًا)، كما قد يختلف هذا المتوسط حسب متغير النوع ومتغير الحالة الاجتماعية ومتغير السن (معًا)، وبالتالي يزداد عدد المجموعات كلما زاد عدد المتغيرات المستقلة، وهذا يقتضي أن تكون العينة كبيرة بما يضمن وجود العدد الكافي من المفردات في كل مجموعة فرعية وفي الوقت نفسه يضمن عدم الإخلال بشرط أساسي في استخدام المقارنات، ألا وهو التقارب بين عدد المفردات في المجموعات التي تتم المقارنة بينها.

– «المجموعات المستقلة» هي المجموعات المتمايزة في الصفة أو الخاصية المقاسة، فالذكور مجموعة متمايزة والإناث مجموعة متمايزة؛ أي أننا بصدد مجموعتين مستقلتين، وعندما نصنف عينة من المبحوثين حسب محل الإقامة إلى (ريف، حضر، بدو)، فإننا نكون أمام مجموعات مستقلة أو متمايزة.

– «المجموعات غير المستقلة» يقصد بها المجموعات المرتبطة، بمعنى أن المجموعة الأولى هي نفسها المجموعة الثانية، أو متطابقة معها، ويتضح ذلك على سبيل المثال في تطبيق المقياس على مجموعة من المبحوثين ثم إعادة تطبيقه مرة أخرى على المجموعة نفسها، ففي هذه الحالة نكون بصدد مجموعتين مرتبطتين (غير مستقلتين)، كما أن المجموعات المرتبطة (غير المستقلة) تكون أيضًا في حالة أن تكون مفردات العينة عبارة عن أزواج متناظرة أو متطابقة (Matched Pairs). وهنا يتعين الكشف عن اتجاه وحجم الفروق بينها.

– تستخدم الأساليب الإحصائية اللابارامترية في التحقق من الفروض الفارقة ذات المتغيرات الاسمية والمتغيرات الرتبية، وتختلف هذه الأساليب حسب عدد المجموعات (العينات) وطبيعتها (مستقلة أو غير مستقلة)، وكذلك حسب طبيعة المتغيرات (اسمية أو رتبية) على النحو التالي:

(أ) قياس الفروق بين المجموعات من حيث المتغيرات الاسمية:

- الفروق لعينة واحدة: اختبار العشوائية، اختبار ذو الحدين، كا^2، اختبار كولموجروف سمير نوف.

- الفروق بين مجموعتين مستقلتين: اختبار فيشر (Fisher Test)، اختبار كا^2 (Chi Square Test) (x^2) اختبار الوسيط (The Median Test).

- الفروق بين مجموعتين غير مستقلتين: اختبار ماكنمارا (McNamara Test).

- الفـــروق بـــين ثـــلاث مجموعـــات مســـتقلة (أو أكـــثر): اختبـــار كـــا^2 للاســـتقلالية (Chi-Square Test for Independence)، اختبار الوسيط (The Median Test).

- الفروق بين ثلاث مجموعات (أو أكثر) غير مستقلة: اختبار كوكران (The Cochran Q Test).

(ب) قياس الفروق بين المجموعات من حيث المتغيرات الرتبية:

- الفروق بين مجموعتين مستقلتين: اختبار كولموجروف - سـمير نـوف (Kolmogrov-Smirnov Test)، اختبار مان- وتني (Mann-Whitney) (U Test) ، وهـذا الاختبـار يسـتخدم كبـديل عـن اختبـار (T) عندما لا توافر شروط اختبار (T) أو عندما يصعب تطبيق اختبار (T) لسبب أو لآخر.

- الفروق بين مجموعتين غير مسـتقلتين: اختبـار ولكوكسـون (Wilcoxon Test)، اختبـار الإشـارة (Sign Test).

- الفروق بين ثلاث مجموعات (أو أكثر) مستقلة: اختبـار كروسـكال- والـيس (Kruskal- Wallis Test)، وهو بديل لابارامتري لتحليل التباين ذي الاتجاه الواحد (One-Way ANOVA).

- الفروق بين ثلاث مجموعات (أو أكثر) غير مستقلة: اختبار فريدمان (The Friedman Test).

هذا ومن المفيد للباحثين متابعة التطور في طرق الإحصاء اللابـارامتري خاصـة الـبرامج الإحصـائية والكتيبات الإلكترونية (Electronic Manuals) والدوريات المتخصصة.

مصادر الفصل الرابع ومراجعه

(أ) مصادر ومراجع عربية:

- رجاء محمود أبو علام (1999)، مناهج البحث في العلوم النفسية والتربوية، (القاهرة: دار النشر للجامعات).

- زكريا الشربيني (1990)، الإحصاء اللابارامتري في العلوم النفسية والاجتماعية والتربوية، (القاهرة: مكتبة الأنجلو المصرية).

- سعدي شاكر حمودي (2000)، علم الإحصاء وتطبيقاته في المجالين التربوي والاجتماعي، (عمان: مكتبة دار الثقافة للنشر والتوزيع).

- سمير كامل عاشور& سامية أبو الفتوح سالم (2003)، العرض والتحليل الإحصائي باستخدام (SPSSWIN)، الجزء الأول، (القاهرة: جامعة القاهرة: معهد البحوث والدراسات الإحصائية).

- سمير كامل عاشور& سامية أبو الفتوح سالم (2005)، العرض والتحليل الإحصائي باستخدام (SPSSWIN)، الجزء الثاني، (القاهرة: جامعة القاهرة: معهد البحوث والدراسات الإحصائية).

- صلاح الدين محمود علام (2000)، تحليل بيانات البحوث النفسية والاجتماعية والتربوية، (القاهرة: دار الفكر العربي)

- عاطف أحمد منصور(1992)، الرياضيات المسلية، (القاهرة: مكتبة ابن سينا).

- عبد الحميد محمد نجم & محمد عبد الهادي المحميد (1990)، الإحصاء الوصفي والتحليلي مع استخدام البرامج الجاهزة، (الكويت: مكتبة جامعة الكويت).

- عزت عبد الحميد محمد حسن (2011)، الإحصاء النفسي والتربوي، (القاهرة: دار الفكر العربي).

- فتحي عبد الله فياض (1991)، التحليل الإحصائي للبيانات الجغرافية، (القاهرة: مكتبة جامعة عين شمس).

- فؤاد أبو حطب & آمال صادق (2010)، مناهج البحث وطرق التحليل الإحصائي في العلوم النفسية والاجتماعية والتربوية، (القاهرة: مكتبة الأنجلو المصرية).

- فؤاد البهي السيد (1978)، علم النفس الإحصائي، (القاهرة: دار الفكر العربي).

(ب) مصادر ومراجع أجنبية:

- Creswe, J. W. (1998) Qualitative Inquiry and Research Design: Choosing Among Five Traditions. Thousand Oaks CA: Sage.

- Gerhardt, Kristopher (2004) The Tritone Paradox: An Experimental and Statistical Analysis.; Dissertation Abstracts International: Section B: The Sciences and Engineering, Vol. 64 (7-B), pp. 35-63.

- Gillespie Marie & Tonybee, Janson (2006) Analyzing Media Text. Oxford. Open University Press.

- Hacker, Kenneth L.; Zakahi, Walter R.; Giles, Maury J. (2000) Components of Candidate Images: Statistical Analysis of the Issue-Persona Dichotomy in the Presidential Campaign of 1996 Communication Monographs, Vol. 67 (3), pp. 227-238.

- Huber, John C. (2000) A Statistical Analysis of Special Cases of Creativity.; Journal of Creative Behavior, Vol. 34 (3), pp. 22-203.

- Lee, Sik-Yum; Zhu, Hong-Tu (2000) Statistical Analysis of Nonlinear Structural Equation Models with Continuous and Polychromous Data. British Journal Of Mathematical and Statistical Psychology, Vol. 53 (2), pp. 209-232.

- Morgan, Kai Andito Damali (2002) Boulder and Vail Models: A Statistical Analysis of the Results for the Examination of Professional Practice in Psychology (1988-1996). Dissertation Abstracts International: Section B: The Sciences and Engineering, Vol. 63 (1-B), pp. 541-568.

- Searle, Shayle R. (2002) Statistical Analysis of Data from Designed Experiment. Contemporary Psychology: APA Review of Books, 2002, Vol. 47 (6), 720-722.

الفصل الخامس

تعريف ببعض الأساليب الإحصائية المتعمقة

تمهـــيد

تتعدد الأساليب الإحصائية وتتطور باستمرار، مستفيدة من التطور المتلاحق للبرامج الإحصائية الجاهزة وتطبيقاتها في كافة مجالات المعرفة، وعلى الرغم من أن الكثير من الأساليب الإحصائية المتعمقة حيز الاستخدام منذ بدايات القرن العشرين، إلا أن التطورات التي يتم إدخالها على تلك الأساليب تجعلها ذات أبعاد نظرية وتطبيقية أوسع نطاقًا وأكثر جدوى للبحوث العلمية. في الصفحات القادمة سوف نوضح أمثلة لتلك الأساليب ممثلة في: تحليل التباين المتعدد، تحليل التغاير، تحليل الانحدار، التحليل العاملي،تحليل المسار، التحليل التمييزي، باعتبارها من أكثر الأساليب استخدامًا في العلوم الاجتماعية.

المبحث الأول

تحليل التباين المتعدد

يعتبر تحليل التباين المتعدد (Multivariate Analysis of Variance)، والمعروف اختصارًا (MANOVA) صورة متقدمة من تحليل التباين الأحادي (ANOVA)، ذلك أن تحليل التباين الأحادي يستخدم للمقارنة بين ثلاث مجموعات أو أكثر؛ من حيث متغير تابع واحد فقط (كالمقارنة بين مجموعات الريف والبدو والحضر في الدرجة على مقياس المعرفة السياسية)؛ أي أن هناك متغيرًا تابعًا واحدًا فقط (المعرفة السياسية)، ونريد تقصي

معنوية الفروق بين مجموعات العينة؛ من حيث متوسط الدرجة على الاختبار الذي يقيس هذه المعرفة.

أما التباين المتعدد (MANOVA) فيستخدم لتقصي الاختلافات أو الفروق بين المجموعات في أكثر من متغير تابع، ويشيع استخدام تحليل التباين المتعدد في الدراسات العلمية بمجالاتها المختلفة بما فيها مجال الاتصال، فقد يريد الباحث معرفة الفروق بين مجموعات العينة حسب المستوى الاقتصادي الاجتماعي؛ من حيث : استخدامات الصحف اليومية، استخدامات المجلات الأسبوعية، استخدامات التليفزيون، استخدامات الإنترنت. أي أن لدينا أربعة متغيرات تابعة. وهناك عدة شروط يتعين توافرها حتى يتسنى استخدام (MANOVA)، وأهم هذه الشروط :

- أن يكون عدد المبحوثين في المجموعة الواحدة أكثر من عدد المتغيرات التابعة.
- يتأثر تحليل التباين متعدد المتغيرات بالقيم المتطرفة، ولذلك يتعين التأكد من أنه لا توجد قيم متطرفة لأحد المبحوثين، سواء في أيٍّ من المتغيرات التابعة كل على حدة، أو في هذه المتغيرات مجتمعة (إذا وجدت حالات ذات قيم متطرفة يتم استبعادها)، وهناك إجراءات معينة في هذا الشأن؛ حيث يتم استخدام الانحدار الخطي للمتغيرات، ومقارنة قيمة (Mahalanobis) بقيمة حرجة يتم الحصول عليها من جداول كا^2 للقيمة الحرجة، وإذا زادت قيمة درجة المبحوث عن هذه القيمة الحرجة، فإن هذا المبحوث يعتبر حالة متطرفة.
- درجات المبحوثين في مجمل المتغيرات التابعة يفضل أن تتخذ شكل التوزيع الطبيعي (يمكن التغاضي عن هذا الشرط إذا كان عدد المفردات في كل خلية من خلايا الجدول لا يقل عن 30).

- وجود علاقة ارتباط خطية بين كل زوج من المتغيرات التابعة (يمكن التأكد من ذلك بعمل تخطيط انتشار بين كل زوج من أزواج المتغيرات التابعة).

- وجود ارتباط متوسط أو معقول بين المتغيرات التابعة بعضها وبعض، فإذا تبين وجود ارتباطات ضعيفة يمكن استخدام تحليل التباين الأحادي، أما عندما تكون الارتباطات قوية (0.8 فأكثر)، فإن ذلك ينتج عنه ما يعرف بالمصاحبة الخطية المتعددة (Multicollinearity)، ويمكن التغلب على ذلك إما بحذف أحد أزواج المتغيرات التابعة ذات الارتباط القوي، أو دمج هذا الزوج ليصبح متغيرًا واحدًا.

وبموجب تطبيق التباين المتعدد (MANOVA) تتنوع المعطيات الإحصائية ذات القيمة العالية، فمن خلال المعطيات الإحصائية المعنونة (Estimated Marginal Means) يمكن معرفة أي مجموعات العينة حصلت على متوسط أعلى أو أقل من المجموعات الأخرى. كما يمكن معرفة ما إذا كانت هناك فروق جوهرية بين مجموعات العينة من حيث التجميع الخطي للمتغيرات التابعة، فإذا كانت دلالة (Wilks' Lambda) أقل من 0.05 تكون هناك فروق جوهرية بين المجموعات، أما إذا كانت هذه الدلالة تساوي أو أكبر من 0.05، فإن ذلك يعني عدم وجود فروق جوهرية بين المجموعات.

وهناك أيضًا الدلالة الإحصائية لمعامل (Hoteling' Trace) وكذلك معامل (Pillai's Trace)) كما تتيح المعطيات الإحصائية الناتجة عن تطبيق تحليل التباين المتعدد التعرف على ما إذا كانت هناك فروق جوهرية بين المجموعات؛ من حيث أي متغير من المتغيرات التابعة، كل متغير على حدة، وذلك من المعطيات الإحصائية المعنونة(Tests of Between Subjects Effect) ، كما يمكن معرفة نسبة التباين في المتغير التابع التي يمكن تفسيرها بالمتغير المستقل (وهو الذي صنفت على أساسه مجموعات العينة)، ويعبر عن ذلك في المعطيات الإحصائية بقيمة تربيع إيتا الجزئي (Partial Eta Square) فإذا كانت هذه القيمة مثلًا 0.22، فإن ذلك يعني أن المتغير المستقل (المتغير الذي صنفت على أساسه مجموعات العينة) يفسر 22% من التباين في المتغير التابع.

المبحث الثاني

تحليل التغاير

يمكن تعريف تحليل التغاير (Covariance Analysis) (ANCOVA) بأنه أسلوب إحصائي للتحكم في المتغيرات التي تتداخل مع المتغير المستقل في التأثير على المتغير التابع. هذه المتغيرات التي يتم ضبطها أو التحكم فيها تسمى بالمتغيرات المصاحبة أو الملازمة (Covariates)، ويعتبر تحليل التغاير من أبرز الأساليب الإحصائية التي يشيع استخدامها في العلوم الاجتماعية،فهو من أساليب الضبط الإحصائي الهامة، كما أنه في بعض جوانبه امتداد للمعاملات الإحصائية التي تتقصى الفروق بين عدد من المجموعات؛ من حيث متوسط الدرجة على مقياس معين مثل اختبار (T) وتحليل التباين الأحادي (ANOVA)، فهو يجمع بين هذه المعاملات وتحليل الانحدار.

وفي الدراسات التي تستخدم تحليل التباين، قد يرى الباحث ضرورة استخدام تحليل التغاير بهدف زيادة قوة اختبار تحليل التباين الأحادي (ANOVA)، بل إن الكثير من الدراسات العلمية الجادة لا تتوقف عند نتائج تحليل التباين، وتذهب إلى ضرورة فهم التباين الذي لايمكن التنبؤ به أو توقعه في نموذج تحليل التباين، وهو تباين الخطأ (التباين داخل المجموعات)، وقد يؤدي ذلك إلى نتائج أكثر دقة. ويدل التباين داخل المجموعات على الانحراف غير المضبوط (والذي يعزى إلى العشوائية في التصميمات التجريبية الكاملة)، ففي تحليل التباين (ANOVA) من المعروف أن تباين الخطأ (أو متوسط مربعات داخل المجموعات) هو باقي أو خطأ التقدير؛ لأن المتغير المستقل يكون هو المنبئ الوحيد.

وفي ضوء الإطار النظري والرؤية الموضوعية المدققة قد يتوقع الباحث أن المتغير التابع يمكن أن يتأثر بمتغير آخر أو بمتغيرات أخرى غير المتغير المستقل أو المتغيرات المستقلة والتي تم تحديدها، (فالباحث الذي يهتم بمعرفة ما إذا كان مستوى المعرفة السياسية

يختلف حسب مستوى التعرض للمضمون السياسي في وسائل الإعلام، قد يتوقع أن مستوى المعرفة السياسية يمكن أن يتأثر بمتغيرات أخرى مثل المشاركة في الأنشطة السياسية، التحدث في الموضوعات السياسية مع الآخرين...)؛ أي أن متغير المعرفة السياسية (كمتغير تابع) يمكن أن تختلف قيمته حسب متغيرات أخرى (وليس فقط حسب مستوى التعرض للمضمون السياسي في وسائل الإعلام)، وتكون دراسة هذه المتغيرات الأخرى من متطلبات زيادة كفاءة التنبؤ بالمتغير التابع وتخفيض تباين الخطأ.

ويتم تطبيق تحليل التغاير في البحوث التجريبية وشبه التجريبية على نطاق واسع؛ إذ إن التصميم التجريبي الجيد هو ذلك الذي يضمن أن يكون التأثير الواقع على المتغير التابع ناتجًا عن المتغير المستقل، مع التحكم في المتغيرات والشروط والظروف التي تنعكس على هذا التأثير. ويمكن استخدام تحليل التغاير أحادي الاتجاه، وتحليل التغاير ثنائي الاتجاه. إن تحليل التغاير أحادي الاتجاه يعني وجود متغير واحد مستقل ومتغير واحد تابع، أما تحليل التغاير ثنائي الاتجاه فيعني وجود متغيرين مستقلين، ومتغير واحد تابع، وفي كل الأحوال يتم التحكم في المتغيرات المستقلة الأخرى.

ويفترض تحليل التغاير أن تكون العلاقة بين المتغير المصاحب / المتغيرات المصاحبة والمتغير التابع / المتغيرات التابعة هي علاقة خطية (Linear)، وفي حالة استخدام أكثر من متغير مصاحب، فإن تحليل التغاير يفترض أن تكون العلاقة خطية أيضًا بين كل زوج من هذه المتغيرات، وعلى مستوى كل مجموعة على حدة - من مجموعات المتغير المستقل- فإن تحليل التغاير يفترض تجانس ميل الانحدار، بمعنى أن تكون العلاقة بين المتغير المصاحب والمتغير التابع على مستوى المجموعة الأولى متشابهة مع العلاقة بين المتغير المصاحب والمتغير التابع على مستوى المجموعة الثانية، وكذلك على مستوى المجموعة الثالثة.....، ومعنى متشابهة أن تكون متقاربة في القيمة، وفي حالة استخدام أكثر من متغير مصاحب، يتعين فحص الارتباط بين هذه المتغيرات بعضها ببعض، وأن يكون الارتباط بينها ضعيفًا،

وإذا تبين أن الارتباط بينها قويّا أو متوسطًا، يفضل حذف أحد هذه المتغيرات، خاصة المتغير الذي يرتبط ارتباطًا قويّا ببقية المتغيرات المصاحبة.

لتوضيح أهمية تحليل التغاير، نفرض أن لدينا ثلاث مجموعات حسب مستوى الدخل (منخفض، متوسط، مرتفع)، وقد كشف تحليل التباين أحادي الاتجاه (NOVA) عن أنه لا توجد فروق جوهرية بين هذه المجموعات الثلاث في الدرجة على مقياس استخدام وسائل الإعلام، فالمجموعة الأولى ذات الدخل المنخفض، لا تختلف عن المجموعة الثانية (ذات الدخل المتوسط)، ولا عن المجموعة الثالثة (ذات الدخل المرتفع)، كما أن المجموعة الثانية لا تختلف عن المجموعة الثالثة ($p > 0.05$)، هنا نجد أن المتغير المستقل هو (مستوى الدخل)، أما المتغير التابع، فهو (استخدام وسائل الإعلام). غير أن الباحث لم يكتف بهذه النتيجة، لأنه يريد معالجة إحصائية أعلى للبيانات الخاصة بالنتيجة السابقة، وبالتالي قام بتقصي العلاقة بين استخدام وسائل الإعلام ومتغيرات أخرى، من بينها المستوى التعليمي فتبين له أن هناك ارتباطًا جوهريّا بين كل من المستوى التعليمي ومستوى استخدام وسائل الإعلام، في هذه الحالة يتم إجراء تحليل التغاير؛ بحيث يكون «مستوى التعليم» هو المتغير المصاحب (Covariate).

مثال توضيحي:

في دراسة تطبيقية كان الهدف معرفة ما إذا كانت معلومات الجمهور عن الوضع في جنوب السودان، تختلف باختلاف مستوى التعرض للمواد السياسية التي تنشرها وسائل الإعلام. اختار الباحث عينة عشوائية قوامها 400 مفردة روعي فيها الضوابط المطلوبة، وتم تطبيق استبيان يتضمن مقياسًا لمستوى التعرض للمواد السياسية التي تنشرها وسائل الإعلام، ومقياسًا للمعرفة بالوضع في جنوب السودان وقت إجراء الدراسة، بالإضافة إلى مقياس ثالث لقياس مستوى الاهتمام بالوضع في جنوب السودان (حيث كان من المتوقع أن تزداد المعرفة بزيادة الاهتمام). عولجت البيانات إحصائيّا، وتم تصنيف المبحوثين إلى ثلاث مجموعات حسب مستوى التعرض للمواد السياسية التي تنشرها وسائل الإعلام (منخفض، متوسط،

مرتفع). كان التساؤل هو : هل توجد فروق جوهرية بين المجموعات الثلاث من حيث المعرفة بالوضع في جنوب السودان؟ باستخدام تحليل التباين أحادي الاتجاه (ANOVA)، خلصت الدراسة إلى النتيجة الموضحة بالجدول الآتي:

ANOVA

Knwlede

	Sum of Squares	df	Mean Square	F	Sig.
Between Grpups	18.025	2	9.013	.128	.880
Within Groups	28020.452	397	70.580		
Total	28038.478	399			

واضح من هذا الجدول أنه لا توجد فروق جوهرية بين المجموعات الثلاث؛ من حيث المعرفة بالوضع في جنوب السودان($F= 0.128; p > 0.05$)، ومعنى ذلك أن هذه المعرفة لا تختلف باختلاف مستوى التعرض للمواد السياسية التي تنشرها وسائل الإعلام حسبما كشف عن ذلك تحليل التباين الأحادي. أراد الباحث أن يتقصى تأثير متغير «الاهتمام بالوضع في جنوب السودان» كمتغير مصاحب (Covariate Variable)، وتم تنفيذ تحليل التغاير، فأسفر عن النتيجة المبينة بالجدول الآتي:

Tests of Between-Subjects Effects

Dependent Variable: Knowlede

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	10986.924[a]	3	3662.308	85.052	.000
Intercept	32613.387	1	32613.387	757.403	.000
Interest	10968.899	1	10968.899	254.738	.000
MediaExposure	346.071	2	173.036	4.019	.019
Error	17051.553	396	43.059		
Total	453599.000	400			
Corrected Total	28038.478	399			

a. R Squared = .392 (Adjusted R Squared = .387)

يكشف هذا الجدول عن أن قيمة (F) والمبينة قرين المتغير (Media Exposure) بلغت 4.019 وهذه القيمة ذات دلالة إحصائية بمستوى معنوية 0.019 ($p < 0.05$)، ومعنى ذلك أنه بعد ضبط متغير«مستوى الاهتمام بالموضوع» - تبين أن درجة معرفة المبحوثين بالوضع في جنوب السودان تختلف اختلافًا جوهريًا باختلاف مستوى التعرض للبرامج السياسية التي تعرضها وسائل الإعلام، هذه النتيجة الجديدة لم يكن من الممكن التوصل إليها بدون استخدام تحليل التغاير.

وننوه هنا إلى أن تحليل التباين (ANOVA) قد يكشف عن وجود اختلافات جوهرية بين مجموعات عينة البحث ($p < 0.05$)، من حيث متوسط الدرجة على مقياس معين، لكن تحليل التغاير يكشف عن أنه لا توجد اختلافات جوهرية ($p > 0.05$)، كما أن تحليل التغاير قد يؤكد نتائج تحليل التباين، سواء فيما يتعلق بوجود أو عدم وجود فروق جوهرية بين مجموعات عينة البحث في الدرجة على مقياس معين، وهنا أيضًا تكون للنتيجة قيمة علمية عالية، فهناك بحوث تفترض وجود تأثير لمتغير مصاحب وأخرى تفترض عدم وجود مثل هذا التأثير.

المبحث الثالث

تحليل الانحـــدار

أولًا: الفكرة الأساسية للانحدار

الانحدار (Regression) في معناه العام هو أسلوب إحصائي يمكن بواسطته التنبؤ بقيمة متغير معين بمعلومية قيمة متغير آخر بناءً على وجود ارتباط جوهري بين المتغيرين. وقد سمي معامل الانحدار بهذا الاسم؛ لأن المعادلة الخاصة به تنحو في تقديراتها للقيم المختلفة نحو المتوسط؛ أي تنحدر نحو المتوسط، فإذا افترضنا أن هناك ارتباطًا جوهريًا بين متغيرين: المتغير الأول (س) وهو عدد الساعات التي يقضيها المفحوصون يوميًا في قراءة الصحف، والمتغير الثاني (ص) وهو عدد الساعات التي يقضونها يوميًا في مشاهدة

التليفزيون، فإننا يمكننا التنبؤ بالوقت الذي يقضيه أي مفحوص في مشاهدة التليفزيون، إذا علمنا الوقت الذي يقضيه في قراءة الصحف، كما يمكننا معرفة الوقت الذي يقضيه في قراءة الصحف بمعلومية الوقت الذي يقضيه في مشاهدة التليفزيون.

ولتبسيط الفكرة الأساسية في معامل الانحدار، افترض أن الصحفيين العاملين في إحدى الصحف اليومية يتقاضى كل منهم مرتبًا شهريًا ثابتًا قدره (ألف دولار أمريكي)، بالإضافة إلى (مائة دولار) عن كل إعلان يجلبه أي صحفي للصحيفة، في هذه الحالة، فإن مجمل الدخل الشهري سيتغير حسب عدد الإعلانات، بمعنى أن عدد الإعلانات التي يجلبها الصحفي، يتغير معه مجمل الدخل أيضًا بمقدار معين:

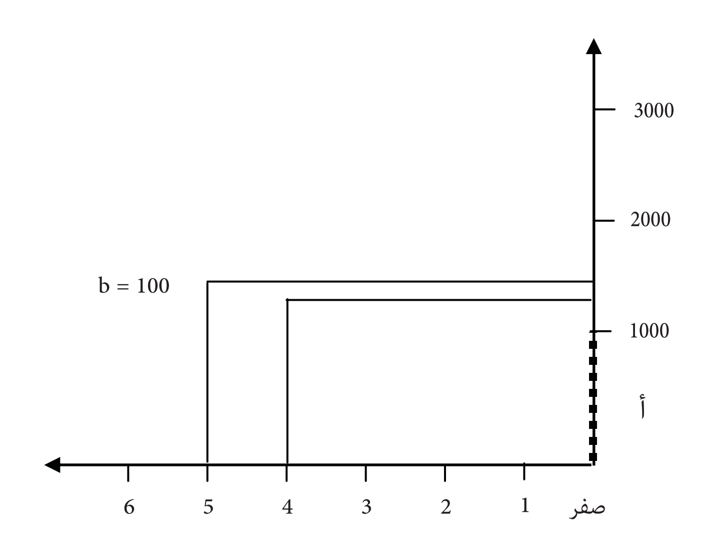
فإذا كان أحد الصحفيين قد حصل على خمسة إعلانات، فإن مجمل دخله يكون: 1000+ (5×100) = 1500

وإذا كان صحفي آخر قد حصل على سبعة إعلانات، فإن مجمل دخله يكون: 1000+ (7×100) = 1700

وإذا كان صحفي ثالث قد حصل على عشرين إعلانًا، فإن مجمل دخله يكون: 1000+ (20×100) = 3000

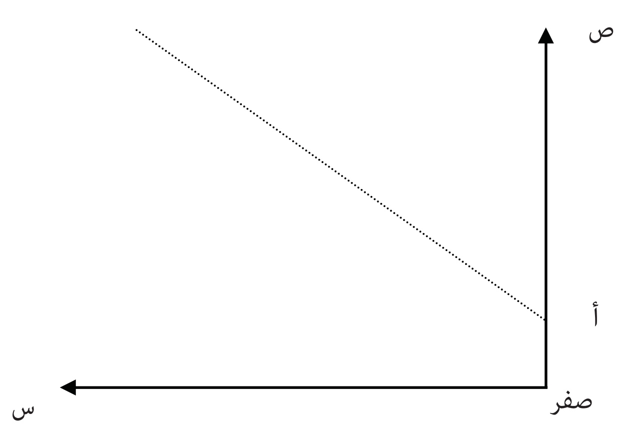
أما إذا كان أحد الصحفيين لم يحصل على أي إعلان، فإن مجمل دخله الشهري يكون: 1000+ (صفر×100) = 1000

من الواضح أن ذلك يعكس وجود ارتباط موجب (Positive Correlation) بين مجمل الدخل وعدد الإعلانات؛ أي أنه كلما زاد عدد الإعلانات، زاد مجمل الدخل فيما يمثله الشكل التالي:



ومع أخذ قيمة الثابت في الاعتبار، والتعبير عن قيمة التغير في المتغيرين، فإن الصيغة المعبرة عن ذلك هي معادلة الانحدار، إنها ببساطة تعبر عن حجم التغير في المتغير التابع لكل وحدة تغير في المتغير المستقل، وبذلك يمكننا التنبؤ بقيمة المتغير التابع إذا علمنا قيمة معينة للمتغير المستقل، ويمكن إيجاد معادلة خط الانحدار باستخدام الدرجات الخام أو باستخدام الدرجات المعيارية أو باستخدام طريقة الانحرافات أو باستخدام معامل الارتباط، ومن معادلة الانحدار يمكننا معرفة قيم المتغير (ص) من واقع قيم المتغير (س) والعكس.

ويقوم الانحدار على نفس الأسس التي يقوم عليها معامل ارتباط بيرسون، خاصة فيما يتعلق بوجود علاقة خطية بين متغيرين من النوع الكمي أو تكويدهما تكويدًا كميًا، فإذا كانت العلاقة موجبة خطية (Linear) بين المتغيرين، فإنها تأخذ الشكل التالي:



إن الخط المنقوط يسمى خط حسن المطابقة (line of best fit) وهو يشير إلى علاقة خطية مستقيمة بين المتغيرين، فجميع النقاط على خط واحد مستقيم، ويعرف بخط الانحدار، وهو ذلك الخط الذي يجعل مجموع الانحرافات عنه أقل ما يمكن، سواء فيما يخص المتغير (س) أو فيما يخص المتغير (ص)، ويمكن من خلاله تحديد أي نقطة على المحور الصادي، إذا علمنا النقطة المقابلة لها على المحور السيني، والعكس صحيح (بمعنى أنه يمكن تحديد أي نقطة على المحور السيني، إذا علمنا النقطة المقابلة لها على المحور الصادي)، فإذا رمزنا للمتغير المستقل بالرمز X وإلى المتغير التابع بالرمز Y، فإن معادلة الانحدار هي:

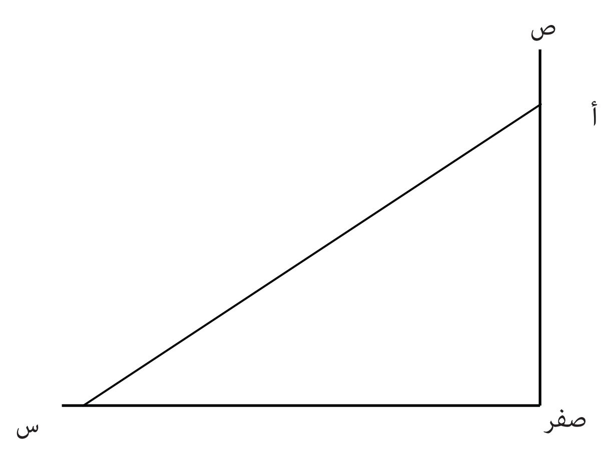
$$Y= a + bx + e$$

فالرمز Y هو المتغير التابع والرمز X هو المتغير المستقل، أما الرمز a فهو قيمة ثابتة، وتسمى الثابت (Intercept) وهي النقطة التي يكون عندها الخط المستقيم (line of Best fit) قاطعًا للمحور الرأسي، أما الرمز b، فيعني ميل هذا الخط، هذا الميل (Slope) هو معامل الانحدار، ويعني المعدل الذي يتغير بموجبه المتغير المستقل بما يستتبعه تغير في المغير التابع، فإذا تغير المتغير المستقل بمقدار معين، فإن المتغير التابع هو الآخر يتغير بمقدار معين، بينما يشير الرمز e إلى الخطأ (Error) ويعني التباين في قيمة المتغير التابع،

وهذا التباين يفسر بعوامل أخرى وليس بالمتغير المستقل، وسوف نتجاهل هذا الرمز بهدف التبسيط. فإذا اعتبرنا أن (س) هو المتغير المستقل، وأن (ص) هو المتغير التابع، فإن معادلة الانحدار عبارة عن تقدير قيمة المتغير التابع استنادًا إلى قيمة الثابت + (قيمة المتغير المستقل× معامل الانحدار)؛ أي أن:

ص= الثابت + (س× قيمة الانحدار)

لكن معامل الارتباط بين المتغيرين قد يكون سالبًا (Negative Correlation)، وفي هذه الحالة نضع إشارة ناقص (-) بدلًا من إشارة زائد (+)، وعندما يكون الارتباط بين المتغيرين سالبًا، فإن هذا الارتباط يمثله الشكل التالي:



وهنا تكون معادلة الانحدار هي:

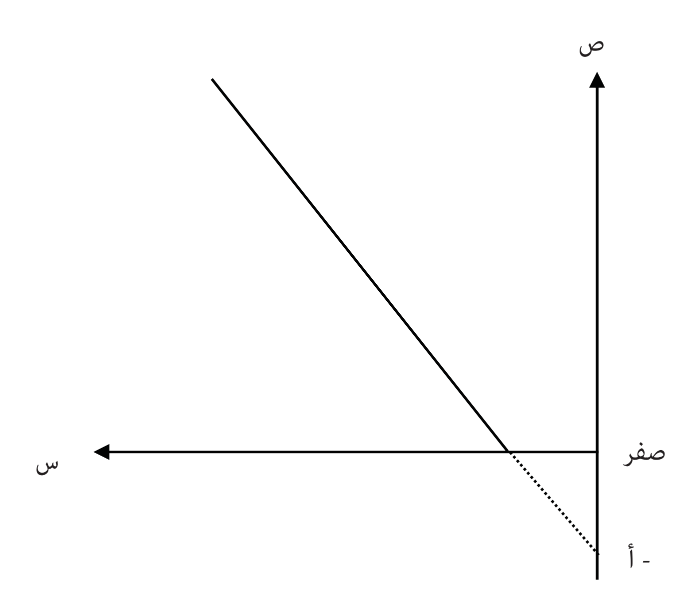
ص= الثابت - (س× معامل الانحدار)

لنفرض أن معادلة الانحدار التي استخرجها النموذج هي:

ص= 50 – 2 س، فإذا كانت قيمة س تساوي 12، في هذه الحالة تكون:

ص= 50 – (2 × 12) = 26

وفي بعض الحالات يكون الخط المستقيم يميل إلى أن يكون رأسيًا متقاطعًا مع المحور الأفقي، في هذه الحالة فإن الثابت يكون بالناقص (Minus) كما هو في الشكل الآتي:



وهنا تكون معادلة الانحدار كالآتي:

ص= - الثابت + س × معامل الانحدار

فإذا كان الثابت يساوي (- 7)، وكانت قيمة (س) هي 23 ومعامل الانحدار 3، فإن

ص= - 7 + (3×23) = 62

ويغطي تحليل الانحدار مدى واسعًا من التطبيقات والاستخدامات بما يخرج عن نطاق أهداف هذا الفصل، من هنا سيتم التعريف بالجوانب الأساسية للانحدار من خلال نماذج مبسطة تتمثل في الانحدار الخطي البسيط، الانحدار المتعدد (مع الإشارة إلى الانحدار غير الخطي)، ثم الانحدار اللوجستي.

ثانيًا: الانحدار الخطي البسيط

الانحدار الخطي البسيط هو التطبيق البسيط أو المباشر للفكرة الأساسية التي يقوم عليها الانحدار. فهو يهدف إلى التنبؤ بقيمة متغير معين بمعلومية متغير آخر؛ أي أنه يتضمن متغيرين فقط، أحدهما هو المتغير المستقل والثاني هو المتغير التابع. مثال ذلك تطبيق الانحدار الخطي البسيط على الدخل الشهري من جهة، والمبلغ المالي الذي يتم إنفاقه على شراء الصحف اليومية من جهة ثانية. إن الدخل الشهري يعامل كمتغير مستقل (Independent)، أما المبلغ النقدي الذي يتم إنفاقه على شراء الصحف اليومية فيعامل كمتغير تابع (Dependent)، غير أنه يمكن استخدام تحليل الانحدار البسيط للتنبؤ بقيمة المتغير (س)، بمعلومية قيمة المتغير (ص)، وكذلك التنبؤ بقيمة المتغير (ص) بمعلومية قيمة المتغير (س).

لنفرض أن دراسة علمية أجريت على عينة عشوائية من المفحوصين لمعرفة العلاقة بين استخدام وسائل الإعلام والمعرفة السياسية، ولنرمز إلى استخدام وسائل الإعلام بالرمز (س)، وإلى المعرفة السياسية بالرمز (ص)، فإذا أردنا التنبؤ بدرجة المعرفة السياسية (ص) بمعلومية درجة استخدام وسائل الإعلام (س)، سمي هذا النوع من التنبؤ بانحدار (ص) على (س)، أما إذا أردنا التنبؤ بدرجة استخدام وسائل الإعلام (س) من درجة المعرفة السياسية (ص)، سمي هذا النوع انحدار (س) على (ص)، وذلك على النحو التالي:

(أ) انحدار (ص) على (س):

تأخذ معادلة انحدار (ص) على (س) الصورة الجبرية الآتية :

$$\text{ص} = \text{ر} \times \frac{\text{ع ص}}{\text{ع س}} (\text{س} - \text{م س}) + \text{م ص}$$

حيث إن :

ص = الدرجة المجهولة وهي درجة المعرفة السياسية التي نريد معرفتها من خلال الدرجة (س) المعروفة لدينا، وهي درجة استخدام وسائل الإعلام.

س = الدرجة المعلومة (درجة استخدام وسائل الإعلام)

ر = معامل الارتباط بين الدرجة(س) والدرجة (ص)

ع س = الانحراف المعياري لمتوسط الدرجة (س)

ع ص = الانحراف المعياري لمتوسط الدرجة (ص)

م س = متوسط درجة الاختبار (س)

م ص = متوسط درجات الاختبار (ص)

لنفرض أن قيم هذه المعطيات فيما يخص معدل استخدام وسائل الإعلام (س) والمعرفة السياسية (ص)، كانت على النحو الآتي:

ر = 0.90

ع ص = 2.6

ع س = 7.56

م س = 10

م ص = 8

بموجب هذه المعطيات يمكن حساب انحدار (ص) على (س)؛ أي معرفة قيمة (ص) بمعلومية قيمة (س) كالآتي:

ص = 0.90 × $\frac{٢,٦}{٧,٥٦}$ (س- 10) + 8

= 0.31 (س-10) + 8

= 0.31 س – 3.1 + 8

= 0.31 س + 4.9

إذن، ص = 0.31 س +4.9

وهذه هي معادلة انحدار (ص) على (س) أو معادلة التنبؤ المطلوبة، فإذا كانت قيمة (س) تساوي 3، فإن قيمة ص = 0.31× 3 + 4.9 = 5.83 تقريبًا

ويمكن الاستفادة من ذلك في رصد أي درجة للمتغير (ص) من واقع معرفة أي درجة للمتغير (س)، فإذا كانت قيمة (س) تساوي 6 مثلًا، فإن قيمة (ص) تساوي:

6.76= 4.9 +6 × 0.31

وفي التطبيق العملي، فإن استخدام الحزمة الإحصائية للعلوم الاجتماعية (SPSS) أو غيرها من البرامج الإحصائية الجاهزة، يتيح كافة المعطيات الإحصائية بما فيها المعطيات اللازمة لتكوين معادلة الانحدار، ففي دراسة عن العلاقة بين المعرفة السياسية ومشاهدة التليفزيون، وباستخدام تحليل الانحدار الخطي البسيط، تضمنت المعطيات الإحصائية النموذج التالي:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.920	.026		35.462	.000
	TV. Viewing	.321	.016	.428	20.454	.000

a. Dependent Variable: political Knowledge

فهذه المعطيات تتضمن قيمة الثابت (constant)، وهذه القيمة هي 0.920 كما تتضمن قيمة الانحدار، وهذه القيمة تساوي 0.428 وبالتالي فإن معادلة الانحدار تكون:

المعرفة السياسية = 0.920 + 0.428 مشاهدة التليفزيون

فإذا كانت قيمة مشاهدة التليفزيون = 3 فإن المعرفة السياسية

= 0.920 + (0.428 × 3) = 2.204

(ب) انحدار (س) على (ص):

كما سبقت الإشارة، فإن انحدار (س) على (ص)، يعني التنبؤ بقيمة المتغير (س) بناء على معرفة قيمة (ص)، ويمكن الحصول على معادلة انحدار (س) على (ص)، بالطريقة نفسها، مع إحلال (س) محل (ص) حيث تتمثل هذه المعادلة في:

س = ر × (ع س / ع ص) (ص- م ص) + م س

بالاستفادة من المعطيات السابقة، عن المتغيرين (س)، (ص) وبالتعويض في تلك المعادلة، نجد أن:

س = 0.90 × (٧,٥٦ / ٢,٦) (ص- 8) + 10

= 2.62 (ص-8) + 10

=2.62 ص - 20.96 + 10

= 2.62 ص - 10.96

أي أن قيمة (س) تساوي 2.62ص – 10.96

فعندما تكون قيمة ص = 6، فإنه يمكن التنبؤ بقيمة (س) كالآتي:

س = 2.62 × 6- 10.96 = 4.76

ولما كان المتغير (س) هو الدرجة على مقياس استخدام وسائل الإعلام، والمتغير (ص) هو الدرجة في المعرفة السياسية، فإن الشخص الذي حصل على 6 درجات في المعرفة السياسية، سوف يحصل على 4.76 في الدرجة على مقياس استخدام وسائل الإعلام.

ثالثًا: الانحدار المتعدد

إذا كان الانحدار الخطي البسيط يتعامل مع متغيرين فقط، فإن الانحدار المتعدد (Multi Regression) يتعامل مع أكثر من متغيرين، فقد يكون لدينا أكثر من متغير مستقل، متغيران مستقلان أو ثلاثة متغيرات مستقلة أو أكثر، ونريد معرفة تأثيرها في متغير تابع، هنا يكون الانحدار المتعدد امتدادًا للانحدار الخطي البسيط، كل ما في الأمر أن الباحث في حالة استخدام الانحدار المتعدد يتعامل مع أكثر من متغير مستقل(More Than Independent Variable). فإذا افترضنا أن لدينا متغيرًا تابعًا (ص)، ومتغير مستقل أول (س1)، ومتغير مستقل آخر (س2)، وأن العلاقة بين المتغير التابع والمتغيرين المستقلين هي علاقة خطية مستقيمة، فمعنى ذلك أن:

ص= الثابت + س1 (ح1)+ س2 (ح2)

فإذا كان الرمز (س1) يعني المتغير المستقل الأول، والرمز (س2) يعني المتغير المستقل الثاني، فإن ح1 تعني قيمة انحدار المتغير المستقل الأول، كما أن ح2 تعني قيمة انحدار المتغير المستقل الثاني، والصيغة الإنجليزية، للمعادلة المذكورة هي:

$$Y = a + b1x1 + b2x2$$

حيث a تعني قيمة الثابت، *b1* تعني المتغير المستقل الأول، *b2* تعني المتغير المستقل الثاني، أما *x1* فتعني انحدار المتغير المستقل الأول، *x2* تعني انحدار المتغير المستقل

الثاني.كمثال لتوضيح هذه الفكرة نفرض أن دراسة علمية تناولت العلاقة بين المعرفة بالقضايا العربية (ص) وكلٍّ من:قراءة الصحف اليومية (س1) ومتابعة الأخبار في التليفزيون (س2)؛ أي أن لدينا متغيراً تابعًا (ص) و متغيرين مستقلين (س1، س2)، وأن هناك ارتباطًا خطيّا موجبًا بين المتغير التابع والمتغيرين المستقلين، كما تبين من المعالجة الإحصائية للبيانات أن الثابت = 5.8، وأن انحدار المتغير المستقل الأول هو 0.311 بينما بلغ انحدار المتغير المستقل الثاني 0.089 في هذه الحالة فإن معادلة الانحدار تكون:

المعرفة السياسية = 5.8 + 0.311 قراءة الصحف اليومية+ 0.089 متابعة الأخبار في التليفزيون.

فمعامل انحدار المتغير المستقل الأول (قراءة الصحف اليومية) هو تقدير (Estimation) لحجم التغير في المعرفة السياسية إذا تغيرت قراءة الصحف بمقدار وحدة واحدة، فإذا تغيرت قراءة الصحف بمقدار وحدة واحدة، فإن المعرفة بالقضايا العربية تتغير بمقدار 0.31 وهكذا فيما يخص متابعة الأخبار في التليفزيون، فإذا تغيرت هذه المتابعة بمقدار وحدة واحدة، فإن المعرفة بالقضايا العربية تتغير بمقدار0.09

بالإضافة إلى ذلك، فإن الانحدار المتعدد إذا كان يوضح لنا مقدار التغير الذي يساهم به كل متغير مستقل على حدة في المتغير التابع، فإن هذا المقدار هو متفرد (Unique) بمعنى أنه يخص هذا المتغير المستقل فقط؛ أي مساهمة هذا المتغير المستقل دون سواه في إحداث التغير في المتغير التابع، كما يكشف لنا تحليل الانحدار المتعدد عن مجمل التغير الذي تحدثه جميع المتغيرات المستقلة (مجتمعة) في المتغير التابع.

وعند إجراء الانحدار المتعدد يجب التأكد أولًا من أن الارتباط بين كل متغيرين من هذه المتغيرات المستقلة ليس قويّا (فلا يتجاوز 0.8)، لأن الارتباط القوي بين المتغيرات المستقلة ينتج عنه ما يعرف بالمصاحبة الخطية المتعددة (Multicollinearity)، بما يؤدي إلى

تداخل تأثير المتغيرات المستقلة في المتغير التابع، فيكون حجم التغير في المتغير التابع بفعل المتغير المستقل الأول متداخلًا مع حجم التغير في المتغير التابع بفعل المتغير المستقل الثاني أو الثالث، ويصعب ضبط المصاحبة الخطية المتعددة إذا كان لدينا أكثر من متغيرين.

من جهة أخرى، فإن قيمة معامل الانحدار الخاص بمتغير مستقل معين لا يدل في حد ذاته على أهمية هذا المتغير في المتغير التابع، فإذا كان انحدار المتغير المستقل (س1) أكبر من انحدار المتغير المستقل الآخر (س2)، فهذا لا يعني أن المتغير (س1) أشد تأثيرًا في المتغير التابع- مقارنة بالمتغير (س2)، فقد تكون وحدات كل متغير مختلفة عن وحدات المتغير الآخر، فالدخل والتعليم مثلًا متغيران مستقلان قد يؤثران في متغير تابع معين، غير أن وحدة قياس الدخل مختلفة تمامًا عن وحدة قياس التعليم، ومن هنا - ومن خلال تحليل الانحدار- يتم الحصول على وزن بيتا (Beta Weight) أو معامل الانحدار المقنن (Standardized Beta Coefficient) الذي يعتمد على الدرجات المعيارية، وهذا هو الأساس السليم الذي نعتمد عليه في المقارنة بين المتغيرات المستقلة من حيث مساهمة كل منها في التغيير الذي يحدث في المتغير التابع؛ إذ إن معامل الانحدار المقنن يدل على عدد الوحدات المعيارية التي يتغير بها المتغير التابع إذا تغير المتغير المستقل؛ بوحدة معيارية واحدة.

وهناك اعتبارات أخرى هامة تتصل بهذه النقطة، وهي العلاقة بين الانحدار المقنن والانحدار غير المقنن، لنفرض أن هناك دراسة أجريت على عينة عشوائية من العاملين واهتمت الدراسة بتحليل العلاقة بين الرضا الوظيفي (كمتغير تابع) وبين كلٍّ من: روتين العمل، الاستقلالية، السن، الدخل.. باعتبارها متغيرات مستقلة ترتبط ارتباطًا خطيًا بالرضا الوظيفي. في هذه الحالة. وبموجب المعالجة الإحصائية للبيانات المستمدة من الدراسة المذكورة (دراسة الرضا الوظيفي)، نفرض أننا حصلنا على المعطيات الإحصائية الآتية:

المتغيرات المستقلة	معامل الانحدار	معامل الانحدار المقنن
الاستقلالية	0.573	0.483
الدخل	0.00124	0.383
روتين العمل	0.168-	0.217-
الثابت	1.936-	-

من هذه المعطيات نتبين أن متغير السن (Age) غير متضمن في الجدول، وهـذا راجـع إلى أن متغـير السن تم استبعاده تلقائيّا من نمـوذج الانحـدار، لأنـه لم يقابـل المعـايير المطلوبـة إحصـائيّا، لقـد أبقـى النموذج على المتغيرات ذات التفسير الجوهري في المتغير التابع (الرضا الوظيفي)، وبناء عـلى المعطيـات الإحصائية الموضحة في الجدول، فإن معادلة الانحدار تكون:

الرضا الوظيفي= -1.936 +0.573 الاستقلالية + 0.00124 الدخل -0.168 روتين العمل.

فإذا كان أحد المفحوصين قد حصل على 16 درجة في مقيـاس الاسـتقلالية، وأن دخلـه 8000 دولار، كما أنه حصل على 8 درجات على مقياس روتين العمل.. هنا تكون درجته في الرضا الوظيفي :

الرضا الوظيفي= - 1.936 + (16×0.573) + (8000×0.00124)-(8×0.168) = 15.81

ومن المعطيات الإحصائية في الجدول السابق نتبين أن متغير (روتين العمل) قد حقق قيمة سالبة (-0.168)، وهذا يعني أن زيادة الروتين تقترن بانخفاض الرضا الوظيفي، كما نتبين أن متغير الاستقلالية قد حقق أكبر قيمة انحدار مقنن وأكبر قيمة انحدار غير مقنن، وإذا كانت قيمة الانحدار المقـنن لمتغـير(الاسـتقلالية) تبلـغ 0.483 فـإن ذلـك يعنـي أنـه

إذا زاد متغير الاستقلالية بمقدار وحدة واحدة، فإن الرضا الوظيفي يزيد بمقدار 0.483 انحراف معياري مع استبعاد تأثير كل من الدخل وروتين العمل. أما فيما يخص متغير الدخل فنجد أنه يقدم أقل قيمة انحدار غير مقنن، بينما يقدم ثاني أكبر قيمة انحدار مقنن، وهذا يعني أن من الخطأ الاعتماد على الانحدار غير المقنن في تحديد أهمية المتغيرات المستقلة من حيث تأثيرها في المتغير التابع، لأن الانحدار غير المقنن تتأثر قيمته بمدى الدرجة الخام على المقياس المستخدم (فقد تكون الدرجة على مقياس معين تتراوح مابين صفر إلى 50، بينما تتراوح الدرجة على مقياس آخر ما بين 4 إلى 10 مثلًا)، ومن هنا فإن الانحدار المقنن (Standardized Regression) هو الذي نعتمد عليه في تقرير أهمية المتغيرات المستقلة في التأثير على معدل التغير في المتغير التابع، فكلما ارتفعت قيمة الانحدار المقنن الخاص بمتغير مستقل، دل ذلك على أن هذا المتغير (المستقل) يحدث تغييرًا أكبر في المتغير التابع.

ويعتبر معامل التحديد (Coefficient Of Determination) عاملًا أساسيًا في الانحدار، ويرمز لهذا المعامل بالرمز ($ر^2$) وبالإنجليزية (R^2)، وتنحصر قيمة معامل التحديد بين الصفر والواحد الصحيح، وهو يعبر عن مجمل تأثير المتغيرات المستقلة في المتغير التابع، فإذا كانت قيمة معامل التحديد هي 0.76، فإن ذلك يعني أن المتغيرات المستقلة التي يتضمنها نموذج الانحدار تفسر 76% من التباين في المتغير التابع، أما النسبة الباقية (24%)، فإنها تفسر بعوامل أخرى، أما معامل التحديد المصحح (The Adjusted Coefficient of Determination) فهو قيمة معامل التحديد بعد الأخذ بالاعتبار عدد المتغيرات المستقلة وعدد مفردات العينة، ويكون معامل التحديد المصحح أقل في قيمته عن معامل التحديد. (في حالة الانحدار الخطي البسيط، فإن قيمة معامل التحديد تساوي مربع قيمة ارتباط بيرسون، فإذا كان ارتباط بيرسون هو 0.8 فإن معامل التحديد يكون 0.64).

الجزئية الأخري ذات الأهمية في معادلة الانحدار المتعدد تتعلق بالخطأ المعياري للتقدير (The Standard Error of The Estimation) والذي يظهر دائمًا ضمن المعطيات الإحصائية. من خلال الخطأ المعياري للتقدير يمكننا الحصول على حدود الثقة في معادلة الانحدار المتعدد. فإذا كان معامل الانحدار لمتغير الاستقلالية هو 0.573 والخطأ المعياري هو 0.09 وكانت عينة الدراسة عشوائية، فإننا يمكننا تقدير الانحدار في المجتمع بدرجة ثقة معينة، فإذا اتخذنا درجة الثقة 95%، فإن حدود الثقة في التقدير تكون ما بين:

0.573- (1.96× 0.09) إلى 0.573+ (1.96× 0.09)

(أي ما بين 40% إلى 75%)

أما إذا اتخذنا درجة الثقة 99%، فإن حدود الثقة في التقدير تكون ما بين:

0.573- (2.58× 0.09) إلى 0.573+ (2.58× 0.09)

(أي ما بين 34% إلى 81%)

وهناك افتراضات يقوم عليها تحليل الانحدار الخطي سواء البسيط أو المتعدد، فهو يتطلب أن تكون العينة كبيرة وأن تكون عشوائية (Random)، وأن تكون المتغيرات من النوع الفتري. غير أنه يمكن استخدام المتغيرات الاسمية في معادلة الانحدار المتعدد، فيمكن مثلًا التنبؤ بقيمة متغير تابع كمي من النوع الفتري (Interval) لمتغير مستقل اسمي (Nominal) مثال ذلك التنبؤ بالدرجة على مقياس استخدام التليفزيون حسب متغير النوع (ذكور & إناث)، أو التنبؤ بدرجة الرضا الوظيفي حسب متغير جهة العمل (صحيفة يومية، قناة تليفزيون، محطة إذاعة). في هذه الحالة يتم تكويد المتغير المطلوب التنبؤ به، ويسمى المتغير الرمزي (Dummy Variables) بمعنى ترميز أقسام المتغير الاسمي بقيم كمية، كأن نرمز إلى الذكور بالرمز (1) وإلى الإناث بالرمز (صفر)، وفي حالة وجود أكثر من مجموعتين في المتغير الاسمي- كمتغير جهة العمل مثلًا، صحيفة يومية،

قناة تليفزيون، محطة إذاعة- نقوم بترميز ثنائي لكل مجموعة، بمعنى تكويد العاملين في الصحيفة اليومية بالرقم (1)، وغير العاملين في تلك الصحيفة بالرقم (صفر)، وهنا تكون الصحيفة (كجهة عمل) بمثابة مجموعة أولى، فالكود (1) يدل على العاملين في الصحيفة، أما الكود (صفر) فيدل على غير العاملين في الصحيفة. المنطق نفسه فيما يخص العاملين في قناة التليفزيون (كمجموعة ثانية)، والعاملين في المحطة الإذاعية (كمجموعة ثالثة)؛ أي أن أقسام المتغير الاسمي تصبح مجموعات (متغيرات) ثنائية التصنيف؛ بحيث يدل الرقم (1) على الانتماء للمجموعة،بينما يدل الرقم (صفر) على عدم الانتماء لهذه المجموعة، ويتم إجراء الانحدار الخطي على هذا الأساس.

بين الانحدار الخطي والانحدار غير الخطي:

عند استخدام الانحدار يتعين أولًا رسم الشكل الانتشاري (Scatter Diagram) للمتغيرين لمعرفة ما إذا كانت العلاقة بينهما خطية أو منحنية، وإذا كان الانحدار الخطي البسيط يستخدم في حالة وجود علاقة خطية/ مستقيمة بين المتغيرين، إلا أنه في الكثير من البحوث لا تكون العلاقة بين المتغيرين علاقة خطية (Linear)، وإنما تكون علاقة منحنية أو غير خطية (Nonlinear)، وفي هذه الحالة يتعين إيجاد أفضل منحنى مطابقة، أو أفضل دالة رياضية تطابق البيانات، سواء كانت الدالة الأسية (Exponential Function) أو دالة القوة (Power Function) أو الدالة اللوغاريتمية (Logarithmic Function) أو دالة القطع المتكافئ (Parabola Function) وعند حساب العلاقة بين المتغير الاسمي والمتغير الفتري نستخدم نسبة الارتباط (η) – وتنطق إيتا. وتستند نسبة الارتباط إلى الخفض النسبي لخطأ التنبؤ بقيمة متغير فتري بمعلومية أقسام متغير نوعي (اسمي)، ويقتضي ذلك وجود مستويين أو أكثر للمتغير المنبئ (Predictor Variable) وهو في هذه الحالة المتغير الاسمي، وذلك لكل قيمة من قيم المتغير الذي نريد أن نتنبأ بقيمته، والذي يكون من النوع الفتري.

رابعًا: الانحدار اللوجستي

في هذا النمط من التحليل تستخدم كلمة لوجستي (Logistic) كصفة مشتقة من كلمة (Log) وهو مختصر اللوغاريتمات (Logarithms) واللوغاريتم ببساطة هو «الأس» الذي يجب أن يرفع إليه الأساس ليعطي عددًا معينًا، وكمثال على ذلك، فإن:

7^3 تنطق هكذا (7 أس 3) وهي تساوي 343؛ أي أن 7^3 (7 أس 3) =343 بمعنى أن الرقم 7 مضروبًا في نفسه ثلاث مرات (7×7×7) يساوي 343، ويسمى العدد 7 بالأساس، أما الرقم 3 فيسمى «بالأس» أي أن العدد 7 يجب أن يرفع إلى الأس 3 ليعطي العدد 343، هذا «الأس» الذي يجب أن يرفع إليه العدد 7 ليعطي العدد 343 هو المعروف باللوغاريتم. وعند تحويل 7^3 (أو 7 أس 3) من الصورة الأسية إلى الصورة اللوغاريتمية يصبح:

7 لو 343=3

وتقرأ كالآتي: لوغاريتم 343 للأساس 7 يساوي 3

وهناك اللوغاريتمات الطبيعية التي أساسها الرمز (E)، وعلى الرغم من أنه لا توجد قيمة محددة له تمامًا إلا أن قيمته التقريبية 2.7182818000 ؛ أي 2.718 علمًا بأن هناك اللوغاريتمات العشرية التي يتخذ فيها الرقم 10 كأساس لها كما هو معروف.

ويرتبط الانحدار اللوجستي ارتباطًا وثيقًا بالاحتمالات، فإذا افترضنا أن احتمال قراءة صحيفة يساوي 70%، فإن هذا يعني أن احتمال عدم قراءة تلك الصحيفة هو 30% وهنا تكون أرجحية قراءة الصحيفة:

$$= 0.7 \div 0.3 = 2.33$$

أما أرجحية عدم قراءة الصحيفة فهي:

= 0.3 ÷ 0.7= 43,

وإذا كان الانحدار الخطي (Linear Regression) يبحث معدل التغير في المتغير التابع Y عندما يتغير المتغير المستقل X بمقدار معين، فإن الانحدار اللوجستي (Logistic Regression) يبحث في معدل التغير في لوغاريتم الأرجحية (Log Odds) لحدوث المتغير التابع Y عندما يتغير المتغير المستقل X.

إن نسبة الأرجحية هي احتمال وقوع الحدث مقسومًا على احتمال عدم وقوعه، فإذا كان احتمال مشاهدة التلفاز 0.68، فإن احتمال عدم المشاهدة هو 0.32 وتكون نسبة الأرجحية هي 0.68 ÷ 0.32 = 2.125، ويتم التعرف على دلالة قيمة تلك النسبة من خلال «اختبار والد» (Wald Test) فإذا كانت قيمة هذا الاختبار معنوية أو دالة (Significant)، فإن هذا يعني أن المتغير المستقل له تأثير جوهري في ترجيح احتمال وقوع أو عدم وقوع الحدث، قد تكون هذه الدلالة عند مستوى المعنوية 0.05 ($p < 0.05$) أو عند مستوى المعنوية 0.01 ($p < 0.01$).

ويتم استخدام الانحدار اللوجستي (Logistic Regression) عادة عندما يكون المتغير التابع ثنائيّا متقطعًا (Binary Dichotomous) بمعنى أن هذا المتغير له حالتان فقط مثل:

- مشاهدة التليفزيون مقابل عدم مشاهدة التليفزيون.
- قراءة الصحف مقابل عدم قراء الصحف.
- سماع الإذاعة مقابل عدم سماع الإذاعة.
- استخدام الإنترنت مقابل عدم استخدام الإنترنت.

ويمكن إعادة تكويد المتغير المستمر (غير المتقطع) بحيث يصبح متغيرًا ثنائيًا متقطعًا، لنفترض أن نتائج أحد البحوث أفادت أن مشاهدة التليفزيون من واقع عينة قوامها 300 مفردة (ن= 300) جاءت كالآتي:

مشاهدة التليفزيون	ك	%
لا يشاهد	75	25
أقل من ساعة	100	33.3
من ساعة إلى ساعتين	80	26.7
أكثر من ساعتين	45	15
المجموع	300	100

في هذه الحالة يمكن إعادة تكويد استجابات المفحوصين؛ بحيث تضم فئتين فقط :

- يشاهد: العدد 225

- لا يشاهد: العدد 75

فالمتغير المتصل (المستمر) - وهو عدد ساعات مشاهدة التليفزيون يوميًا- تم تحويله إلى متغير متقطع ثنائي (له حالتان فقط)، وبموجب هذا التحويل يمكن تطبيق نموذج الانحدار اللوجستي، وعادة ما يتم تكويد احتمال الحدوث بالرقم (1) أما احتمال عدم الحدوث، فيتم تكويده بالرقم (صفر) فإذا افترضنا أن حدوث التصويت في الانتخابات يأخذ الرقم (1)، فإن عدم حدوثه يأخذ الرقم (صفر)، وهذا التكويد (Coding) يتم أثناء إدخال البيانات باستخدام أحد البرامج الإحصائية الحاسوبية المعروفة.

كمثال توضيحي،نفرض أن المتغير التابع المطلوب التنبؤ بحدوثه هو «التصويت في الانتخابات البرلمانية»، وعند معالجة البيانات فإن هذا المتغير يعامل كمتغير تابع (Dependent variable)، ومن الواضح أنه من النوع الثنائي (Dichotomous)،

فالمفحوص إما أنه سوف يدلي بصوته أو أنه لـن يـدلي بصـوته في الانتخابـات. أمـا مجموعـة المتغـيرات المستقلة (Independent) أو المنبئة لنفرض أنها تتمثل في:

- مشاهدة التليفزيون (فيما يخص الانتخابات).
- سماع الإذاعة (فيما يخص الانتخابات).
- قراءة الصحف اليومية الحكومية (فيما يخص الانتخابات).
- قراءة الصحف الحزبية (فيما يخص الانتخابات).
- الانتماء الطبقي.

وإذا كان المتغير التابع (الإدلاء بالصوت في الانتخابات) يتخذ الرقم الكودي (1) بمعنى التصويت في الانتخابات، أو الرقم الكودي (صـفر) بمعنـى عـدم التصـويت في الانتخابـات، فـإن مجموعـة المتغـيرات المستقلة لها قيم كميّة معينة تعكس استجابات المفحوصين حسبما أفادوا بها، هكذا يتم إدخال بيانـات المتغير التابع (الرقم صفر أو الرقم 1)، و إدخال القـيم الكميّـة للمتغـيرات المسـتقلة، ويتعامـل نمـوذج الانحدار اللوجستي مع تلك المتغيرات المستقلة كمتغيرات منبئة (Predicting Variables).

والجدول الآتي يوضح بعض المعطيات الإحصائية لنموذج الانحـدار اللوجسـتي المقـدر باعتبـار أن الموقف من الانتخابات (التصويت أو عدم التصويت) متغير تـابع (Y)، أمـا المتغـيرات المسـتقلة، فهـي : مشاهدة التليفزيون، سماع الإذاعة، قراءة الصحف اليومية الحكومية، قراءة الصحف الحزبية، الانـتماء الطبقي:

النموذج المقدر

	B	S.E	Wald	Df.	Sig.	Exp (B)
مشاهدة التليفزيون	0.471	0.127	13,84	1	0.000	0.624
سماع الراديو	0.472	0.203	5.44	1	0.02	1,604
قراءة الصحف الحكومية	0.492	0.161	9,4	1	0.002	1,636
قراءة الصحف الحزبية	0.592	0.122	18,8	1	0.000	,589
الانتماء الطبقي	0.105	0.139	0.578	1	0.447	0.009
الثابت	0.963	0.377	6.5	1	0.011	2,62

يلاحظ من هذا الجدول أن جميع قيم B موجبة (العمود الأول)، وهذه القيم هي معاملات النموذج المقدر بوحدات اللوغاريتم (Log Odds)، أما العمود الثاني (S.E) فيبين الخطأ المعياري لكل معامل من المعاملات (B)، ويتضمن العمود الثالث معطيات اختبار Wald، وهو في الحقيقة اختبار ($كا^2$) المبني على معامل (Wald) لمعرفة معنوية قيمة كل معامل، ومن الجدول أيضًا وفي ضوء قيمة المعنوية (.Sig) نتبين أن الانتماء الطبقي غير دال إحصائيًا ($Sig. > 0.05$) أما بقية المتغيرات فهي دالة إحصائيًا ($Sig. < 0.05$).

أخيرًا، فإن القيم الموجودة في العمود Exp (B) تعبر عن نسبة الأفضلية لكل معامل، وهذه النسبة هي المعامل مرفوعًا في الأس (E)، وقد سبقت الإشارة إلى أن قيمة هذا الأس 2.718 تقريبا (اللوغاريتم الطبيعي)، ويتم الحصول على معادلة الانحدار اللوجستي بإضافة الثابت (Constant) إلى القيم الواردة في العمود الأول (B)، أما ناتج هذه المعادلة فيرمز له بالحرف (Z)، أو الحرف (ز) وحسب المعطيات الإحصائية بهذا الجدول تكون قيمة (Z) أو (ز) كالآتي:

ز= 0.963+0.105 الانتماء الطبقي + 0.592 قراءة الصحف الحزبية + 0.492 قراءة الصحف الحكومية + 0.472 سماع الراديو + 0.471 مشاهدة التليفزيون.

أما الصياغة الإنجليزية لتلك المعادلة فهي:

Z=0.963+0.105 Class Affiliation+0.592 Reading the Political Parties Newspapers+0.492 Reading the Governmental Newspapers +0.472 Radio Listening +0.472 Television Viewing.

ويستفاد بقيمة ناتج هذه المعادلة أي قيمة (Z) وكذلك بقيمة الأس الطبيعي (E) وقيمته 2.718 في تقدير احتمال حدوث المتغير التابع، وذلك بموجب المعادلة:

$$\text{Pro} = \frac{1}{1 + e(-Z)}$$

إن ناتج تطبيق هـذه المعادلـة يسـمى الاحـتمال المقـدر (The Estimated Probability) أي احـتمال وقوع الحدث، فإذا بلغت قيمة هذا الاحتمال (المقدر) حوالي 0.70 مـثلًا، فـإن ذلـك يعنـي أن أرجحيـة وقوع الحدث تبلغ 0.70 وبالتالي فإن أرجحية عدم وقوعه 0.30 ولما كانـت قيمـة الاحـتمال المقـدر في مثالنا هذا هي 0.70 فإن ذلك يعني أن المتغيرات المستقلة تزيد أرجحية التصويت في الانتخابات بنسبة 0.70، وهذه المتغيرات هي: مشاهدة التليفزيـون، سـماع الإذاعـة، قـراءة الصـحف اليوميـة الحكوميـة، قراءة الصحف الحزبية. وغني عن البيان أن كل العمليات الحسابية يقوم بها الحاسب الآلي.

كما يسفر تطبيق الانحدار اللوجستي عن معطيات إحصائية أخرى هامة مثل: سالب ضعف دالـة الإمكان (2loglikelhood -) حيث توجد قيمة لهذه الدالة في كل خطوة وذلـك لاختبـار كفـاءة النمـوذج، كما تتضمن المعطيات الإحصائية قيمة (R2) (مربـع معامـل الارتبـاط) بطريقـة (Cox & Snell) وكـذلك بطريقة (Nagelkerke)، بالإضـافة إلى مربـع كـاي (Chi Square) ودلالتـه.... ويسـتعان بهـذه المعطيـات الإحصائية في رصد وتفسير النتائج، كما يمكن للباحثين الاستفادة بهـا في إجـراء تحلـيلات إحصـائية أكـثر تطورًا.

أخـيرًا، تجـدر الإشـارة إلى أن تحليـل الانحـدار ذو أهميـة في الدراسـات الإعلاميـة والاجتماعية عمومًا؛ إذ إن هناك العديد من القضايا الحيوية التي يتعين التخطيط لها عـلى

التنبؤ الدقيق، فنحن نريد أن نتنبأ بعمق التفاعل أو التواصل الأسري في ضوء معدل استخدام الإنترنت أو التليفزيون أو غير ذلك من وسائل الإعلام، كما أنه من الأمور الجوهرية للقائمين على حملات ترشيد الاستهلاك، مثلًا التنبؤ بمعدل التعرض للحملة الإعلامية (Media Campaign) في ضوء معدل استخدام وسائل الإعلام، وقد نشط الباحثون في علم الاجتماع بالتنبؤ باتجاه الشباب نحو الجريمة في ضوء معدل التعرض لأفلام ومواد العنف في التليفزيون أو السينما أو الإنترنت، ويستفيد التخطيط الاجتماعي من تلك المعطيات الإحصائية في توعية الأسرة بشأن ضوابط سليمة فيما يخص علاقة الأبناء بوسائل الإعلام. من جهة أخرى، فإن الدراسات الاجتماعية تهتم بالعوامل المؤثرة في الأسرة كوحدة اجتماعية، وكثيرًا ما تنحو هذه الدراسات إلى استخدام أسلوب تحليل الانحدار في التنبؤ بالعوامل ذات التأثير في معدلات الزواج الطلاق والاستفادة من ذلك في التخطيط الاجتماعي.

المبحث الرابع

التحليل العاملي

يتسع نطاق التحليل العاملي نظريًا وتطبيقيًا في واقعه وتطوراته المستمرة، غير أننا في هذه الجزئية سوف نشير فقط إلى الأفكار الأساسية التي من شأنها توضيح أهم جوانب هذا الأسلوب الهام من خلال توضيح أهمية التحليل العاملي والمفاهيم الأساسية له، أنماط وطرق التحليل العاملي، تدوير العوامل وتسميتها وتفسيرها، مع الإشارة إلى أبرز الأخطاء في تطبيق التحليل العاملي، ونشدد هنا على ضرورة الرجوع إلى الكتابات المتخصصة والبرامج الإحصائية الجاهزة.

أولًا: أهمية التحليل العاملي:

يعتبر التحليل العاملي (Factor Analysis) من الأساليب الإحصائية المهمة التي يمكن أن يستخدمها الباحث في تحليل البيانات متعددة المتغيرات، ودراسة العلاقات القائمة بين تلك المتغيرات واختزالها في عدد أقل من العوامل(Factors) التي تفسر الظاهرة مجال الدراسة، فالهدف من استخدام التحليل العاملي يتعلق عادة بتلخيص العلاقات بين المتغيرات بطريقة دقيقة ومنظمة ومقتصدة من أجل فهم أفضل وخلق تصور فكري أوضح للظواهر، فعن طريق التحليل العاملي يمكن اختزال عدد كبير من المتغيرات واستخلاص أكبر قدر من المعلومات منها؛ لأن مجموعة كبيرة من المتغيرات المتعددة يتم تجميعها في عدد قليل من العوامل التي تفسر التباين، مثلما تفسر العلاقات في هذه المجموعة الكبيرة من المتغيرات.

وعلى الرغم من أن التحليل العاملي له بعض الخصائص التي يشترك فيها مع أساليب إحصائية أخرى، إلا أنه يتميز عن هذه الأساليب في كونه يتناول مجموعات كبيرة من البيانات المستمدة من الاختبارات والمقاييس بمختلف أنواعها، ويمكن إجراؤه على عشرات المتغيرات خاصة من خلال البرامج الإحصائية الحاسوبية المتطورة،كما أن التحليل العاملي من الأساليب الإحصائية التي تتميز بالمرونة؛ حيث يمكن استخدامه في تصميمات بحثية متعددة؛ وذلك للتحقق من صحة الفروض، ورسم خرائط المفاهيم، ودراسات الحالات، والدراسات الطولية، كما يمكن الإفادة منه في تحليل بيانات متنوعة، مثل: المسوح الاجتماعية، والسلاسل الزمنية، ودرجات الاختبارات والمقاييس في كافة المجالات.

كما أن التحليل العاملي يسمح بالتصور البصري للعلاقات بين المتغيرات المتعلقة بالظواهر المختلفة عن طريق التمثيل الهندسي، وبناء نماذج مجردة للواقع الاجتماعي

والظواهر المختلفة من خلال المعادلات والتمثيلات الهندسية، فليس غريبًا أن يطلق عليه البعض «حساب العلوم الاجتماعية Calculus of Social Sciences».

ثانيًا: المفاهيم الأساسية في التحليل العاملي

هناك مجموعة من المفاهيم الأساسية التي لا بد أن يعرفها الباحث؛ حتى يحقق فهمًا أفضل للتحليل العاملي، علمًا بأن الاستيعاب الإحصائي لتلك المفاهيم يتطلب المعرفة بأساسيات الرياضيات العالية (خاصة الجبر الخطي وهندسة المتجهات والتفاضل)، وفيما يلي نقدم تعريفًا موجزًا بأهم مفاهيم التحليل العاملي بما يساعد الباحثين على فهم إجراءاته ومعطياته:

(1) العامل:

إن مفهوم العامل (Factor) مفهوم جوهري في التحليل العاملي؛ ولكي نوضح المقصود بهذا المفهوم نقدم المثال التالي:

إذا حللنا العدد 6 إلى عوامله الأولية، فإننا نحصل على المعادلة التالية:

$1\times2\times3 =6$

فالأعداد 3، 2، 1 هي عوامل العدد 6 أو المكونات الرئيسية للعدد 6 فإذا كان الرقم 6 يدل مثلًا على المساحة، فإن العدد 3 يدل على الطول، أما العدد 2 فيدل على العرض، ولا يدل العدد 1 على شيء. لكن عندما يكون الرقم 6 يدل على الحجم فإن العدد 3 يدل على الطول، أما العدد 2 فيدل على العرض، وقد يدل العدد 1 على الارتفاع. في هذا المثال نتبين أن الرقم 6 يعني شيئًا محددًا، وأن الأعداد 3، 2، 1 هي مكونات هذا الشيء، ولابد من تسمية هذه المكونات.

ومن المنظور البحثي، فإن العامل (Factor) هو المعنى العام الـذي يجمـع أكـثر مـن بنـد أو متغـير، ولنتأمل البنود التالية:

- أنا مهتم بالحصول على معلومات عن القضايا السياسية.
- أنا مهتم بمتابعة جلسات البرلمان.
- أميل إلى تأييد المرشحين المعتدلين.
- أشارك في الندوات المعنية بالموضوعات السياسية.
- أتواصل مع المنتديات السياسية على الإنترنت.

هذه البنود الخمسة، ما هو المعنى العام الذي يجمعهـا؟ قـد يكـون المعنـى العـام الـذي يجمـع البنود المذكورة هو (المشاركة السياسية)، وهذا المعنى العام هـو مـا نسـميه العامـل، وهـذه التسـمية يضعها الباحث للعامل (Factor) الذي ترتبط به مجموعة من البنود، بمعنى أنه في ضوء محتـوى البنـود تتم تسمية العامل.

ومن المنظور الإحصائي، فإن العامل يلخص الارتباطات (Correlations) بين البنـود أو المتغـيرات، أي أن العامل (Factor) صيغة رياضية كبديل عـن هـذه البنـود أو تلـك المتغـيرات، وقـد كشـفت دراسات سبيرمان هذه الحقيقـة عنـدما أوضـحت أن «العامـل» هـو السـبب المبـاشر وراء وجـود الارتباط القائم بين أي ظاهرتين، فإذا فرضنا أن الظاهرة (أ) تـرتبط ارتباطًـا موجبًـا بالظـاهرة (ب)، فإن هذا الارتباط سببه وجود عامل مشترك (ش) يؤثر تأثيرًا موجبًا في الظاهرتين (أ) و (ب)، وعندما يختفي تأثير العامل (ش) يتلاشى الارتباط بين هـاتين الظـاهرتين. إن العوامـل (Factors) التـي يهـدف التحليل العاملي الكشفي لاستخلاصها من مصفوفة الارتباطـات بـين المتغـيرات تعـد بمثابـة تكوينـات فرضـية (Hypothetical Constructs) تنطـوي عـلى المتغـيرات، وهنـاك أنـواع مختلفـة مـن العوامـل

لكل منها تباين خاص به، وتتمثل تلك العوامل في: العامل العام، والعوامل المشتركة، والعوامل النوعية.

فالعامل العام (General Factor) هو العامل الذي تتشبع به جميع المتغيرات، فإذا كان لدينا استبانة تتكون من عشرين بندًا لقياس الحاجات التي يشبعها الترفيه لدى الجمهور،فإن العامل العام هو ذلك العامل الذي تتشبع به جميع بنود الاستبانة.

أما العوامل المشتركة (Common Factors)، فكل عامل منها تتشبع به مجموعة من البنود،كأن نتبين أن البنود أرقام:1، 2، 3، 4، 5، 6،7، 12، 18 تتشبع بالعامل المشترك الأول، في حين تتشبع البنود أرقام 8، 9، 10، 11، 19،13،20 بالعامل المشترك الثاني، أما العامل المشترك الثالث، فتتشبع به البنود17،15،16،14.

أما فيما يخص العوامل النوعية (Specific Factors) فهي العوامل التي يتشبع بها بند واحد فقط،فإذا كان هناك أحد العوامل الذي يتشبع به بند واحد، فإن هذا العامل يطلق عليه عامل نوعي.

وتأتي الإشكالية من أن بعض البنود / المتغيرات يمكن أن تتشبع بأكثر من عامل (ولكل عامل مسمى معين)، وهنا يثار التساؤل : إذا كان البند متشبعًا بأكثر من عامل، فإلى أي عامل ينتمي هذا المتغير/ البند؟ والإجابة هي أنه في حالة تشبع المتغير/ البند بأكثر من عامل، فإنه ينتمي للعامل الذي ترتفع قيمة تشبعه، فإذا كان تشبع البند (س) بالعامل الأول هو 0.311 وكان تشبعه بالعامل الثاني هو 0.651، فإن هذا البند يكون منتميًا للعامل الثاني، وحسب دراسات جيلفورد فإن تشبع المتغير بالعامل يكون دالًا إذا كانت قيمة هذا التشبع 0.3 فأكثر.

وفي التحليل العاملي نجد أن العوامل المستخلصة تفسر نسبة معينة من التباين، مثال ذلك أن يسفر التحليل العاملي لبيانات مستمدة من تطبيق مقياس استخدامات وسائل

الإعلام عن خمسة عوامل، وتبلغ قيمة التباين الذي تفسـره تلك العوامل 0.58 وهذا يعني أن العوامل المستخرجة تفسر 58% من التباين، وأن 42% تفسرها متغـيرات أخـرى؛ أي أن العوامـل المستخلصـة لا تفسر التباين بكامله (Total Variance)، وإنما تفسر نسـبة مـن هـذا التبـاين، وتـزداد القيمـة العلميـة للعوامل المستخلصة كلـما زادت نسـبة التبـاين الـذي تـفسره بشرط سـلامة إجـراءات تطبيـق التحليـل العاملي.

(2) مصفوفة الارتباط:

في التحليل العاملي تكون مصفوفة الارتباط (Correlation Matrix) معبرة عن قيم معاملات الارتباط بين المتغيرات بعضها بعضًا، أو العوامل بعضها بعضًا، أو بين العوامل والمتغيرات.

ونقطة البدء في التحليل العاملي هي تكوين مصـفوفة الارتباطـات بـين مجموعـة المتغـيرات التـي عددها (ن)، فإذا كان الاستبيان يقيس أربعة أبعاد هـي: مشـاهدة بـالتليفزيون، سـماع الراديـو، قـراءة الصحف اليومية، قراءة الصحف الأسبوعية، فإن مصفوفة الارتباط تكـون عـلى النحـو المبـين بالجـدول الآتي:

الأبعاد	مشاهدة التليفزيون	سماع الراديو	قراءة الصحف اليومية	قراءة الصحف الأسبوعية
مشاهدة التليفزيون	**1**	0.39	0.58	0.38
سماع الراديو	0.39	**1**	0.35	0.36
قراءة الصحف اليومية	0.58	0.35	**1**	0.53
قراءة الصحف الأسبوعية	0.38	0.36	0.53	**1**

أي أن المتغيرات توضع عـلى كـلٍّ مـن البعـدين الأفقـي والـرأسي، وتوضـع قـيم معامـل الارتباط بين كل متغيرين منها في خلايا المصفوفة. كما يلاحظ أن الجدول منقسم إلى مثلثـين

مشتركين في الوتر، فكل مثلث يتضمن الخلايا نفسها،ويفصل بين المثلثين الخلايا التي بها الرقم (1)، وهذه الخلايا تتضمن معامل ارتباط كل بعد بنفسه، وبذلك تكون مصفوفة الارتباط في المثلث العلوي تشتمل على جميع الارتباطات بين المتغيرات، كما أن هذه الارتباطات نفسها تتضمنها مصفوفة الارتباط في المثلث السفلي؛ لتيسير جمع القيم في أي صف وأي عمود في عمليات استخلاص العوامل.

(3) **الاشتراكيات أو قيم الشيوع:**

تدل الاشتراكيات (Communalities) على التباين المشترك (Common Variance)، ويكون لكل متغير اشتراكية قيمتها مجموع مربعات تشبعات المتغير بكل عامل من العوامل المستخرجة من التحليل، كمثال على ذلك، إذا كان تشبع المتغير (س) بالعامل الأول 0.56 وتشبعه بالعامل الثاني 0.126، فإن قيمة شيوعه = $(0.56)^2 + (0.126)^2$ ؛ أي $0.3136 + 0.015876 = 0.329476$ أي أن قيمة الشيوع فيما يخص المتغير الأول هي 0.33 تقريبًا، وهذه القيمة تشير إلى نسبة إسهام العوامل المشتركة في المتغير، ويرمز لها في الإنجليزية بالرمز h^2، كما تدل مكونات هذه القيمة على إسهام كل عامل على حدة في المتغير، فنسبة إسهام العامل الأول في المتغير المذكور هي 0.3136، أما نسبة إسهام العامل الثاني في هذا المتغير فهي 0.015876

وفي مقابل التباين المشترك، هناك التباين النوعي الخاص بالمتغير (Unique Variance) وقيمته تساوي الواحد الصحيح مطروحًا منه الاشتراكيات، ففي المثال المذكور يكون التباين النوعي للمتغير الأول = 1 - 0.33 = 0.67 تقريبًا

(4) **التشبعات:**

يُقصد بالتشبع (Saturation) للمتغير، ارتباط ذلك المتغير بعامل (Factor) معين تم استخلاصه، فإذا كان المتغير/ البند الأول مثلًا يتشبع بالعامل الخامس بمقدار 0.64، فإن

ذلك يعني أن المتغير/ البند الأول يـرتبط بالعامـل الخـامس بقيمـة 0.64 وتُـفسر قيمـة التشـبع بـنفس طريقة تفسير معامل ارتباط بيرسون؛ حيث يـتم تربيـع قيمـة التشـبع العـاملي لكـلٍّ مـن المتغـيرات في مصـفوفة التشـبعات للحصـول عـلى نسـبة التبـاين في المتغـير التـي يمكـن تفسـيرها بواسـطة العامـل المستخلص. ومتوسط مربعات التشبعات العاملية في هذه المصفوفة يدل على مقـدار التبـاين الكـلي في المتغيرات كمجموعة التي يتم تفسيرها بواسطة هـذا العامـل أو ذاك، أمـا متوسـط مجمـوع مربعـات التشبعات العاملية في جميع الأعمدة، فيدل على نسبة التبـاين الـذي يمكـن تفسـيره بواسـطة العوامـل المستخلصة، وهذا يكون مؤشرًا لمدى صلاحية هذه العوامل في تفسير تباين مجموعة المتغيرات الأصلية.

ويلاحظ أن مجموع مربعات التشبعات في أعمدة مصفوفة التشـبعات العامليـة يسـاوي مجمـوع اشتراكيات المتغيرات، كما أن مجموع نسب التباين في المتغيرات التي يمكـن تفسـيرها بواسـطة العوامـل المستخلصة يسـاوي متوسـط قـيم الاشـتراكيات،وكلما كـان المتغـير مشـتركًا بدرجـة أكـبر مـع غـيره مـن المتغيرات في العوامل المشتركة، زاد مربع قيمة الاشتراكيات. ونظرا لأن عدد العوامل المستخلصــة يكـون أقل من عدد المتغيرات الأصلية، فإننا لا نستطيع الحصـول على قيم المتغـيرات باستخدام العوامل، وإنمـا يمكن تقدير قيمها باستخدام معـادلة الانحدار المتعدد لكل متغير على هذه العوامل.

فإذا افترضنا أن العوامل غير متعامـدة؛ أي غـير مرتبطـة، فـإن مجمـوع مربعـات التشــبعات في أي صـف من صفوف مصفوفة التشبعات العاملية يكون مساويًا مربع معامل الارتباط المتعدد بين العوامل المستخلصة ومتغير معين من مجموعة المتغيرات. ولتوضيح ذلك نفـترض أن لـدينا مصـفوفة التشـبعات العاملية التالية واشتراكياتها التي تم استخلاصها باستخدام أحد أساليب التحليل العاملي، لمقياس يتكون من عشـرة بنود،كما في الجدول التالي:

البنود	العوامل قبل التدوير		
	العامل الأول	العامل الثاني	الاشتراكيات
1-	0.58	0.47	0.5573
2-	0.54	0.43	0.4765
3-	0.46	0.55	0.5141
4-	0.62	0.71	0.8885
5-	0.57	0.61	0.697
6-	0.52	0.59	0.6185
7-	0.66	0.58	0.772
8-	0.71	0.64	0.9137
9-	0.43	0.53	0.4658
10	0.56	0.48	0.544
التباين المشترك	**3.26**	**3.1**	**6.36**
نسبة التباين الكلي	**0.326**	**0.31**	**0.64**
نسبة التباين المشترك	**51.3**	**48.7**	**100**

فالبند الأول مثلًا يرتبط بالعامل الأول بمقدار 0.58، كما يرتبط بالعامل الثاني بمقدار 0.47 أي أن تشبع البند الأول بالعامل الأول يساوي 0.58، أما تشبع البند الأول بالعامل الثاني فهو 0.47 وتدل الأرقام المدونة في خانة الاشتراكيات على نسبة التباين الكلي في البند، هذا التباين الذي يمكن تفسيره بالعاملين الأول والثاني، فهذان العاملان مثلًا يفسران حوالي 56% من تباين استجابات المبحوثين على البند الأول، أما النسبة المتبقية وقدرها 44%، فإنها تفسر بعوامل أخرى وكذلك بأخطاء القياس. ويلاحظ أن التباين المشترك للعامل الأول يساوي 3.26 (وهذه القيمة هي مجموع مربعات تشبعات البنود بالعامل الأول)، أما عن متوسط هذه القيمة فيكون بقسمتها على عدد البنود؛ أي 3.26 ÷ 10= 0.326 وهي نفسها قيمة نسبة التباين الكلي التي تمثل نسبة

التباين التي يفسـرها العامل الأول في البنود العشرة، أما العامل الثاني فيفسر نسبة 0.31 من التبـاين في البنود العشرة، وبالتالي، فإن العاملين الأول والثاني يفسران حوالي 0.64% من التباين الكلي. يلاحظ من الجدول أيضًا أن كل قيمة من قيم الاشتراكيات الخاصة بكل بند تسـاوي مجمـوع مربـع تشـبع المتغـير بالعامل الأول + مجموع مربع تشبع المتغـير بالعامـل الثـاني، خـذ مـثلًا البنـد الأخـير تجـد أن قيمـة الاشتراكيات =0.544 وهذه القيمة تساوي $(0.56)^2 + (0.48)^2$

(5) الجذر الكامن:

الجذر الكامن (Eigenvalue **أو** Latent Root) مصطلح مأخوذ مـن جـبر المصـفوفات، وفي التحليـل العاملي، فإن الجذر الكامن هو حجم التباين الذي يفسـره العامل، وهو يـدل عـلى مجمـوع مربعـات تشبعات هذا العامل بالبنود التي يتشبع بها، وفي هذا التحليـل، فـإن قـيم الجـذور الكامنـة والخاصـة بالعوامل التي يتم استخراجها - تتناقص تدريجيًا، بمعنى أن العامل الأول يكون له أكبر جذر كامن، أمـا العامل الأخير فإنه يكون له أصغر جـذر كـامن، وهـذا يرجـع إلى أن التحليـل العـاملي يسـتخرج الحـد الأقصـى الممكن لتباين كل عامل مع ترتيبها تنازليًا، ويجب أن يكون مجموع الجـذور الكامنـة يسـاوي مجموع الاشتراكيات.

ومن واقع قيمة الجذر الكامن لكل عامل يمكن معرفة نسبة التباين الذي يفسـره هذا العامـل مـن التباين الكلي للعوامل المستخرجة، كما يمكن الاستناد إلى قيمـة الجـذر الكـامن لتحديـد متـى يتوقـف التحليل العاملي (من الشائع - خاصة في طريقة المكونات الأساسية- الإبقاء على العوامـل التـي يكـون جذرها الكامن واحد صحيح فأكثر، بمعنى أن العوامل التي يقل جذرها الكامن عن الواحد الصحيح يتم الاستغناء عنها في التحليل العاملي).

ثالثًا: أنماط التحليل العاملي

هناك نمطان رئيسيان للتحليل العاملي، النمط الأول هو التحليل العاملي الاستكشافي (Exploratory Factor Analysis)، والنمط الثاني هو التحليل العاملي التأكيدي (Confirmatory Factor Analysis) وقد توصل الدكتور / فؤاد أبو حطب إلى هذا التقسيم وتم اعتماده في الأدبيات الأجنبية والعربية المعنية بالتحليل العاملي. إن التحليل العاملي الاستكشافي (Exploratory Factor Analysis) يختص باكتشاف العوامل المثلى التي تتشبع بها بنود المقياس أو الاختبار، مثال ذلك أن يقوم الباحث بإعداد مقياس من خمسين بندًا لقياس استخدامات وسائل الإعلام، ويريد أن يتوصل إلى تصنيف دقيق لهذه البنود. إن التحليل العاملي الاستكشافي قد يسفر عن مجموعة عوامل (ولتكن خمسة عوامل مثلًا) تتشبع بها بنود المقياس، فقد يتضح للباحث أن البنود أرقام : 8، 11، 18، 19، 23 تتشبع بعامل معين، وكذلك كل مجموعة من البنود تتشبع بعامل مختلف.. وهكذا يتمكن الباحث من تحديد الأبعاد (العوامل) التي يتخذها كأبعاد لمقياس استخدام وسائل الإعلام.

أما التحليل العاملي التأكيدي (Confirmatory Factor Analysis) فيقصد به التأكد من تصنيف مسبق لبنود الاختبار أو المقياس، لنفترض أن الباحث استعان بالنظريات والدراسات السابقة في تصميم مقياس الوعي السياسي لدى الشباب، وأن بنود هذا الاستبيان تم توزيعها على عدة أبعاد، فمثلًا البنود من الأول حتى العاشر تقيس الوعي بالقضايا السياسية، أما البنود من الحادي عشر حتى العشرين فتقيس الوعي بأداء البرلمان.. وهكذا تم تصميم المقياس بحيث تختص كل مجموعة من البنود بقياس أحد أبعاد الوعي السياسي. هنا يكون دور التحليل العاملي التأكيدي هو «التأكد» من صحة هذا التصنيف، فإذا كانت البنود من 11 حتى 20 مثلًا قد صاغها الباحث لقياس الوعي بأداء البرلمان، فهل هذه البنود تتشبع بعامل واحد؟ قد يتبين للباحث أن هذه البنود بالفعل

تتشبع بعامل معين، وبالتالي هذا تأكيد لما قام به الباحث، ويكون هذه العامل هو «الوعي بأداء البرلمان»، وقد يتبين للباحث غير ذلك (من واقع العوامل التي تتشبع بها بنود المقياس)، الأمر الذي يقتضي استحداث تصنيف جديد للبنود، أو إعادة صياغة الأسس النظرية. هكذا يتضح أن التحليل العاملي التأكيدي يستهدف اختبار تصنيف على أساسه تتوزع بنود الاختبار أو المقياس.

غير أن التمييز بين الاستخدامين الكشفي والتأكيدي لا يكون دائما تمييزًا قاطعًا؛ لأن الكثير من البحوث تجمع بين الاستخدامين. ومن النادر أن يقوم الباحث بإجراء التحليل العاملي على مجموعة عشوائية من المتغيرات، بل يكون مدركًا بدرجات متفاوتة طبيعة المتغيرات والأبعاد التي تضم هذه المتغيرات.

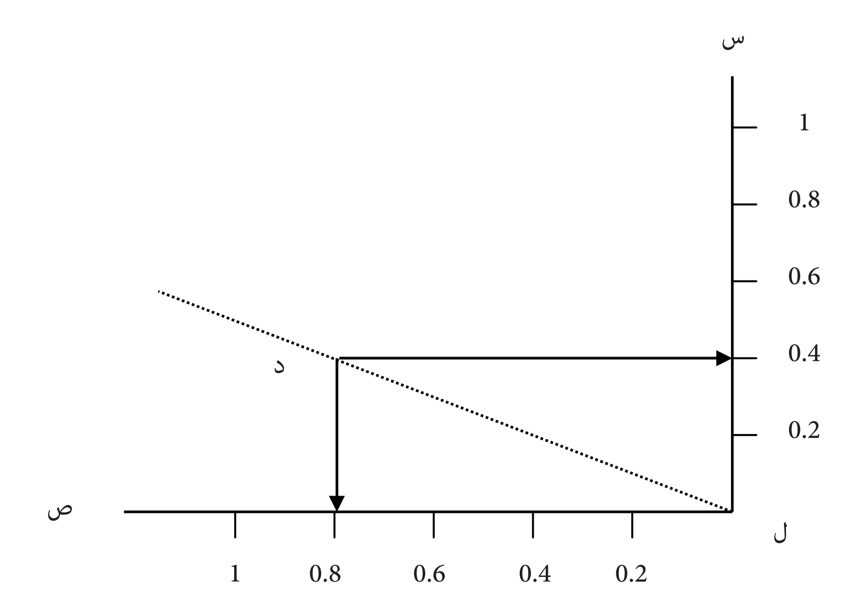
رابعًا: طرق التحليل العاملي

تتعدد طرق التحليل العاملي، كما أنها تتطور باستمرار، ففي البداية كانت الطرق شائعة الاستخدام هي الطريقة القُطرية (Diagonal Method)، والطريقة المركزية (Centroid Method)، وطريقة العوامل المتعددة (Multiple Factors)، لكن هذه الطرق تطورت في أسسها وتطبيقاتها، وكان نتيجة ذلك أن ظهرت طرق حديثة للتحليل العاملي، ومازال هذا المجال يشهد تطورًا ملموسًا خاصة بعد ظهور وتطور البرامج الإحصائية الجاهزة، ومن هذه الطرق : طريقة المكونات الأساسية (Principal Components) وطريقة المحاور الرئيسية (Principal Axis) وطريقة الأرجحية القصوى (Maximum Likelihood) وطريقة ألفا (Alpha factor Analysis) وطريقة الصورة (Image Factor Analysis)، كما ظهرت طرق أخرى أكثر حداثة مثل: طريقة أقل المربعات المرجحة (Un-weighted Least Squares) وطريقة أقل المربعات المعممة (Generalized Least Squares)، ولكلٍّ من هذه الطرق أسسها الرياضية (للاستزادة

حول هذا الموضع ننصح بالرجوع إلى الكتاب الموسوعي للدكتور صلاح الدين علام (2000)، كما أن لهذه الطرق أسبابًا «ترجح» استخدام أيٍّ منها، وإن كانت هذه الأسباب غير قاطعة، ومن واقع مطالعة الدوريات المتخصصة في مجال الاتصال نتبين أن هناك طريقتين هما الأكثر استخدامًا: وهما طريقة المكونات الأساسية، وطريقة المحاور الأساسية. وتقوم طريقة المكونات الأساسية (PC) على تكوين متغيرات جديدة غير مرتبطة، وكل مكونة أو دالة (Function) رئيسية هي تركيب خطي من المتغيرات، كما أن هذه الدوال (Functions) يتم اختيارها بحيث تكون متعامدة. أما طريقة المحاور الرئيسية (PA) فهي قريبة من طريقة المكونات الأساسية، غير أنه في طريقة المحاور الرئيسية يتم وضع القيم التقديرية للاشتراكيات في الخلايا القطرية، وهذه القيم عبارة عن مجموع مربعات تشبعات المتغير (البند) بجميع العوامل، أما في طريقة المكونات الأساسية، فإن هذه الخلايا يوضع فيها الواحد الصحيح. وتمتاز طريقة المحاور الرئيسية بأن كل عامل من العوامل المستخرجة يفسر أكبر قدر ممكن من تباين بنود الاختبار أو المقياس.

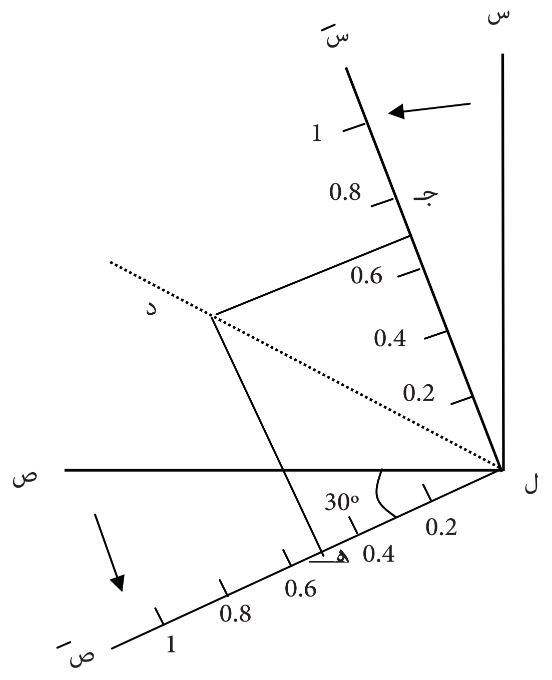
خامسًا: تدوير العوامل

يقصد بتدوير العوامل أن يحل العامل محل المتغيرات (البنود) بحيث تصبح التشبعات بدائل للدرجات المعيارية، أما عن عملية التدوير ذاتها فتعني تدوير العوامل حول نقطة الأصل حتى تصل إلى وضع بديل، كمثال توضيحي لهذه الفكرة نفرض أن أحد المقاييس تم تطبيقه على عينة (ن) وتمت معالجة البيانات عامليًا، وأسفر التحليل العاملي عن أن المقياس يتشبع بعاملين، العامل الأول هو (س) وكان تشبع المقياس على هذا العامل يساوي 0.4 أما العامل الثاني، فهو (ص) وكان تشبع المقياس على هذا العامل يساوي 0.8 (وذلك قبل التدوير). إن التعبير عن ذلك بالرسم يعكس أن العامل (س) والعامل (ص) متعامدان:



يدل الخط (ل د) على متجه المقياس، وتتحدد النقطة (د) في ضوء قيمة تشبع المقياس بكلٍّ من العامل (س) والعامل (ص)، إنها مسافة المتغير على كلٍّ من المحورين كبعدين للمقياس.

فلنتخيل أن هذا الشكل تم تحريك محوريه (ل س)، (ل ص) مع تثبيت نقطة الأصل في موضعها؛ وبحيث تكون حركة المحورين حرة تسمح بتكوين زاوية جديدة قدرها 30 درجة وتصل إلى موضعين جديدين لكلٍّ منهما: الموضع الأول (ل س -)، أما الموضع الثاني فهو (ل ص -)، مع ثبات نقطة الأصل على النحو المبين بالشكل التالي:



في هذا الشكل نلاحظ أن النقطة (د) اتخذت قيما جديدة على تشبع المقياس بالعـاملين (س، ص)، والسؤال هو: ما هي هذه القـيم الجديدة رغم أن موضع العاملين لم يتغير؟، أي ما هي مسافة النقطة (د) على المحور ؟ أليست هذه القيم الجديدة هي مسـافة النقطة (د) على المحـور (ل س-) والمحـور (ل ص-) عند الإسقاط المتعـامد لعمود على كلٍّ من العمودين (د ج)، (د هـ)؟

إن القيم الجديدة يمثلها النقطتان (جـ)، (هـ) على المحور (ل س-) والمحور (ل ص-)، وتبلغ قيمـة الأولى 0.75، بينما تبلغ قيمة الثانية.0.5، وبموجب تدوير العوامل تتم إعادة توزيع نسبة التباين الذي يفسره كل عامل على حدة، وإحداث تكافؤ في أهمية المتغيرات التي يتشبع بها كل بند.

وتحتوي البرامج الإحصائية الجاهزة على عدة أساليب لتدوير العوامل، وتندرج هذه الأساليب تحت أسلوبين: الأسلوب الأول هو التدوير المتعامد (Orthogonal Rotation) والأسلوب الثاني هو التدوير المائل (Oblique Rotation)، ويلاحظ تطور هذه الأساليب باستمرار في النسخ الحديثة للبرامج الإحصائية.

ومن الطرق شائعة الاستخدام في التدوير المتعامد طريقة (Quartimax)، طريقة (Maxplane)، طريقة (Varimax) وهذه الطريقة الأخيرة هي الأكثر استخدامًا لإجراء التدوير العاملي المتعامد، وتؤدي إلى نتائج جيدة، وتعتمد هذه الطريقة على تدوير جميع العوامل (Factors) الممكنة مثنى/ مثنى على حدة إلى أن يتم تعظيم (Maximizing) دالة معينة تتعلق بالتشبعات العاملية.

أما التدوير المائل فيستخدمه الباحثون بهدف استخلاص عوامل مترابطة؛ أي ليست متعامدة كما في التدوير المتعامد؛ حيث إن المحاور المرجعية للعوامل تحصر بينها زوايا ليست قائمة، وقد يتم استخدام التدوير المائل بهدف مقارنة نتائجه بنتائج التدوير المتعامد والمفاضلة بينهما، أو للتوصل إلى استنتاجات معينة، أو ربما يريد الباحث تحديد وتفسير أبسط تكوين ممكن للبنية العاملية والتوصل إلى علاقات سببية.

ومن أبرز طرق التدوير المائل نذكر: طريقة (Promax)، وكذلك طريقة (Oblimax)، وطريقة (Binormamin) وغير ذلك من الطرق، وإن كانت طريقة (Promax) هي الأكثر استخدامًا. وفي هذه الطريقة توضع قيم الاشتراكات في الخلايا القُطرية لمصفوفة الارتباطات بدلًا من الواحد الصحيح، وإجراء التدوير المائل على المصفوفة العاملية الناتجة.

سادسًا: تسمية العوامل وتفسيرها

بعد استخلاص العوامل وتدويرها تدويرًا متعامدًا أو مائلًا، والتوصل إلى مصفوفة التشبعات العاملية بعد التدوير، ينبغي تفسير البنية العاملية الناتجة عن التحليل، وإضفاء

معنى على العوامل بحسب مجال الدراسة، وذلك بتسمية العوامل تسمية مناسبة، والاستعانة بالتمثيل البياني لتوضيحها واستيعابها. وتعتمد تسمية العوامل (Naming Factors) على الهدف من التحليل ومنظور الباحث فيما يتعلق بهذه العوامل، فقد يهدف الباحث- من استخدام التحليل العاملي - توضيح مفاهيم مفيدة في مجال يفتقر إلى المفاهيم الواضحة، وبعض الباحثين يهدف إلى التوصل إلى علاقات علّية (سببية) تفسر أسباب ظاهرة معينة، ففي الحالة الأولى تصف العوامل (Factors) العلاقات القائمة بين المتغيرات، وتصنفها في ضوء التسمية التي يقترحها الباحث، أما في الحالة الثانية فإن الباحث ينظر إلى العوامل كأسباب تنطوي عليها هذه العلاقات، وتكون التسمية متعلقة بهذه الأسباب، وفي بعض الحالات الأخرى ربما لا يحاول الباحث تسمية العوامل أو إضفاء معنى معين عليها، وإنما يرمز لهذه العوامل برموز معينة، ويوضح فقط تشبعاتها ببعض العوامل.

سابعًا: أبرز الأخطاء في تطبيق التحليل العاملي

إن الدراسة العاملية لا بد أن ترتكز على أهداف وفروض / تساؤلات واضحة ومحددة وقابلة للقياس وذات دلالة لموضوع الدراسة وأهدافها وإطارها النظري، كما أن الدراسة العاملية تتطلب التحديد الدقيق للمتغيرات وكفاءة اختيار مفردات العينة والحجم المناسب، وكذلك دقة تصميم أداة / أدوات القياس.

وبعد مرحلة إدخال البيانات يتعين فحص هذه البيانات ومراجعتها واستخراج النسب والتكرارات، والتأكد من أنها سليمة ومتكاملة، وتفحص خصائصها ونزعتها المركزية؛ ومن ثمّ تكوين مصفوفة الارتباطات العاملية، وإذا تبين أن هناك معاملات ارتباط كثيرة غير دالة إحصائيًا أو منخفضة القيمة، فإن ذلك يعني إعادة التفكير في استخدام التحليل العاملي (بالعدول عنه كلية أو إعادة تشفير البيانات).

كما يتعين أن يكون الباحث على وعي تام بشأن ما إذا كان يستخدم التحليل العاملي الاستكشافي أو التأكيدي، وأن يحدد طريقة التحليل العاملي المناسبة، ويقرر أن يحدد عدد العوامل أو لا يحدد هذا العدد، ويختار طريقة التدوير المناسبة وفق معايير علمية، وبموجب تنفيذ هذه الإجراءات والحصول على المعطيات الإحصائية، يكون على الباحث تنظيمها وجدولتها وتفسيرها في ضوء موضوع الدراسة وأهدافها وتساؤلاتها وفروضها وخلفيتها النظرية.

هذه متطلبات أساسية عند تطبيق التحليل العاملي، وإذا كان الإخلال بهذه المتطلبات ينطوي على أخطاء جسيمة، فإنه يعكس سوء فهم هذا التحليل ومتطلباته، فمن مظاهر سوء الفهم أن يقوم الباحث بإجراء التحليل العاملي دون الاستناد إلى تصميم علمي دقيق يأخذ بعين الاعتبار الخطوات المرحلية المطلوبة، وكيفية تنفيذها على النحو السليم، كأن يقوم الباحث بمجرد جمع البيانات وإدخالها في الحاسوب ومعالجتها إحصائيًا دون دراية كافية بمتطلبات إجراء التحليل العاملي ودون وعي بالبرامج الإحصائية الجاهزة وكيفية الاستخدام الأمثل لها،وما تعنيه مخرجاتها وكيفية تنظيم هذه المخرجات والاستفادة بها..كل ذلك يجعل التحليل العاملي للبيانات عملية آلية خاطئة تسفر عن نتائج مضللة أو غير ذات معنى. بالإضافة إلى ذلك هناك مجموعة من الأخطاء الشائعة في التحليل العاملي نخص بالذكر:

(1) عدم التحقق من الافتراضات التي يتطلبها التحليل العاملي في البيانات، والمتعلقة بمستوى قياس المتغيرات وشكل توزيعاتها، فبعض الباحثين يستخدم متغيرات توزيعاتها ملتوية التواء شديدًا أو مبتورة، أو متعددة المنوال، أو مقسمة تقسيمًا ثنائيًا متطرفًا.. لذلك من الضروري فحص توزيع البيانات وخصائصها ونزعتها المركزية، وإعادة تشفيرها إذا تطلب الأمر ذلك، كما يتعين التحقق الإحصائي من عاملية البيانات (Factorability of Data)، وهناك اختبار في هذا الشأن هو

اختبار Barlett، ويتعين إجراء هذا الاختبار للتأكد من أن البيانات صالحة للتحليل العاملي، ويكون ذلك إذا كانت قيمة هذا الاختبار دالة إحصائيًا عند مستوى 0.05 على الأقل ($p < 0.05$).

(2) سوء اختيار عينة الدراسة،فأحيانًا تكون العينة متحيزة / غير ممثلة للمجتمع المستهدف أو غير ذلك من الأخطاء والمثالب في طريقة اختيار العينة. ويتصل بهذه النقطة مسألة حجم العينة، صحيح أن هناك جدلًا بشأن حجم العينة في البحوث العاملية لكن أقل الرؤى تطرفًا في هذا الشأن ترى أن الحد الأدنى لعدد مفردات العينة يتراوح ما بين 100 إلى 150 مفردة، وبالتالي فإن العينة الأقل من ذلك لا تصلح بياناتها للتحليل العاملي، وهناك مَن يرى أن حجم العينة إلى عدد بنود الاستبيان يكون من 10 إلى 1 بمعنى عشر مفردات مقابل البند الواحد (فإذا كان الاستبيان – أو البنود التي ستخضع للتحليل العاملي – 30 بندًا، فيجب ألا يقل حجم العينة عن 300 مفردة، وعلى الرغم من أن بعض الإحصائيين يرى أن التحليل العاملي يمكن إجراؤه على عينات أقل من 100 مفردة، إلا أنهم يؤكدون على ضعف الثقة في النتائج المستمدة من تلك العينات (Prayman &Cramer,1999, p.273)، وهناك اختبارات إحصائية للتأكد من كفاءة العينة في التحليل العاملي، أهمها اختبار(Kaiser-Myer-Olkin) والمعروف اختصارًا (KMO)، وتتراوح قيمة هذا المؤشر مابين الصفر والواحد الصحيح؛ ولكي تكون العينة مناسبة يجب ألا يقل المؤشر عن 0.6 كحد أدنى.

يتصل بالعينة أيضا هدف الدراسة فيما يتعلق بتفسير العوامل، فإذا كان تفسير العوامل يتعلق بالفروق الفردية من حيث سلوك معين لدى فئة عمرية معينة،فإنه ينبغي أن تكون العينة متجانسة نسبيًا في العمر، أما إذا كان تفسير العوامل يتعلق بالفروق العمرية من حيث سلوك معين، فإن العينة ينبغي أن تكون ممثلة لفئات عمرية متباينة، وإذا استخدمت

الدراسة عينة مشتركة من البنين والبنات، يُفضل تحويل درجات كلٍّ منهما إلى درجات معيارية قبل إيجاد قيم معاملات الارتباط، وفي هذه الحالة ينبغي إدخال (النوع) كمتغير في عمليات التحليل، وعزل أثره، ثم إجراء التحليل العاملي على مصفوفة البواقي.

(3) استخدام بيانات تتعلق بمتغيرات متداخلة (غير مستقلة)، كأن يكون أحد المتغيرات مركبًا أو تجميعيّا لمتغيرات أخرى، أو أن تكون الاستجابة على بند معين تعتمد على الاستجابة على بند آخر، وكذلك استخدام متغيرات متداخلة /غير متمايزة من حيث السلوك أو الصفة أو الخاصية محل القياس، ويدخل ضمن هذا الخطأ استخدام صورتين متكافئتين من اختبار أو مقياس معين، ثم معالجة البيانات عامليّا باعتبار الصورتين مقياسًا واحدًا.. كل ذلك ينتج عنه معطيات إحصائية غير دقيقة أو غير ذات معنى، والحصول على عوامل من المستوى الأدنى في التنظيم الهرمي للعوامل (عوامل غير ذات قيمة أو لا معنى لها).

(4) ضعف أو عدم كفاءة المتغيرات /البنود التي تشكل الاستبيان أو الاختبار من حيث شمولها وتمايزها وتركيزها في الموضوع المراد قياسه؛ حيث ينتج عن ذلك قلة عدد العوامل المستخلصة؛ مما يؤدي إلى صعوبات في تدوير العوامل لكي تحقق محكات التكوين البسيط؛ لذلك ينبغي أن يصمم الباحث دراسته؛ بحيث يكون عدد المتغيرات كافيًا لاستخلاص عدد كافٍ من العوامل المتعامدة نسبيّا على الأقل. كما أن عدم الاهتمام الكافي بتصميم الاستبيان ينتج عنه مشكلة أخرى تتعلق بعدد المتغيرات المتشبعة بالعوامل، فالعامل الواحد - لكي يعتد به- يجب أن تتشبع به ثلاثة متغيرات على الأقل، غير أن الباحث قد يفاجأ بأن بعض العوامل تتشبع بها أقل من ثلاثة متغيرات، وهذا يتطلب إعادة النظر في تصميم البحث كلية، مع التدقيق في تصميم أداة جمع البيانات واستيعاب الإطار النظري (أو إعادة النظر فيه) حتى يتسنى للباحث توقع عدد العوامل

المستخلصة وما إذا كانت المتغيرات، ستتوزع توزيعًا منطقيّا / متعادلًا على العوامل.

(5) استخدام أساليب إحصائية غير مناسبة، كأن يستخدم الباحث معاملات ارتباط معينة دون التحقق من أنها تناسب البيانات، أو استخدام قيم اشتراكيات غير مناسبة في الخلايا القُطرية لمصفوفة الارتباطات (كأن يضع الواحد الصحيح في هذه الخلايا عند استخدام التحليل العاملي لمجموعات أو التحليل الطائفي (Group Factor Analysis) فالواحد الصحيح يصلح إذا استخدم الباحث أسلوب المكونات الرئيسة (Principal Components)، ومن مظاهر الاستخدام غير الصحيح للأساليب الإحصائية استخدام طرق غير مناسبة في تدوير المصفوفة العاملية، أو استخدام التدوير المتعامد عندما يكون التدوير المائل أكثر ملاءمة لإعطاء نتائج أفضل، أو عدم الاستناد إلى محكات في عملية التدوير.

(6) استخدام متغيرات/ بنود ليس بسبب أهميتها، وإنما لتوافرها لدى الباحث؛ مما يؤدي إلى تعقيد إجراءات تدوير العوامل وتفسيرها، ومعنى ذلك أن البنود التي يشملها التحليل العاملي يجب اختيارها بعناية، وأن تكون ضرورية للبحث، مع الأخذ في الاعتبار أن عدد المتغيرات يجب أن يزيد عدة مرات على عدد العوامل (فمثلًا ينبغي أن يكن هناك خمسة متغيرات جيدة ومناسبة على الأقل مبدئيّا لكل عامل متوقع استخلاصه)، فإذا كان هناك مقياس يقيس دوافع مشاهدة التليفزيون، وكان الباحث يتوقع أو يفترض- بناء على اعتبارات علمية- أن هذه الدوافع قد تكون طقوسية، وإما أن تكون نفعية، فيجب أن يكون هناك خمسة بنود على الأقل تقيس الدوافع الطقوسية، وخمسة بنود أخرى على الأقل تقيس الدوافع النفعية.

(7) سوء تفسير العوامل: من ذلك مثلًا تفسير العامل الأول الذي يتم استخلاصه على أنه عامل عام، وإعطاء تسمية للعوامل دون فحص طبيعة هذه العوامل وما قد يتطلبه ذلك من إعادة التحقق من صدق التكوين الفرضي لها، وهناك أيضًا خطأ عدم فحص محتوى المتغيرات المتشبعة بها العوامل؛ الأمر الذي يترتب عليه إعطاؤها مسميات غير صحيحة، وهذا يقود إلى استنتاجات مضللة

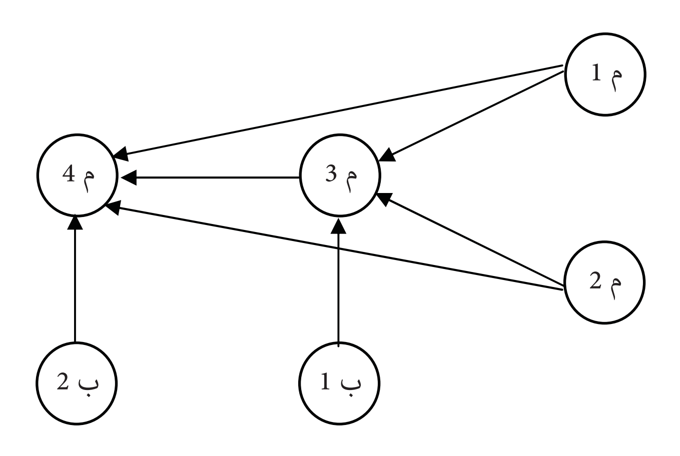
المبحث الخامس

تحليل المسار

تحليل المسار (Path Analysis) هو أسلوب في تحليل البيانات لاختبار صحة نموذج يفترضه الباحث لتفسير العلاقات بين المتغيرات موضع البحث بطريقة متسقة، فإذا افترض الباحث (بناء على أدلة علمية) أن المتغير (س) يؤثر في المتغير (ص) الذي يؤثر بدوره في المتغير (ع) هو نموذج تفسيري لظاهرة معينة، فإن استخدام تحليل المسار يمكن أن يثبت صحة هذا النموذج أو خطأه. أي أن تحليل المسار ليس طريقة للكشف عن العلية أو السببية (Causality)، وإنما هو طريقة لاختبار نموذج سببي (Casual Model) يفترضه الباحث، إنه يقول لنا: «تأثير عدة متغيرات على بعضها البعض»، ومن خلال تحليل المسار يمكن تقدير التأثير النسبي للمتغيرات ضمن هذا النموذج المفترض.

من هنا، فإن تحليل المسار شائع الاستخدام في البحوث العلمية، خاصة في ظل تراكم المعرفة حول الظواهر المختلفة وإمكانية الاستفادة منها في بناء النماذج النظرية التي تفسر هذه الظواهر. ويبدأ تحليل المسار ببناء النموذج النظري الذي يفسر العلاقة بين المتغيرات التي تتم دراستها. وفي بناء مثل هذا النموذج لابد من مراعاة حقيقة أن السبب لا بد أن

يسبق الأثر (The Cause Must Precede the Effect)، وبالتالي لا بد من مراعاة الترتيب الزمني للمتغـيرات التي تحت الدراسة؛ بحيث يظهر ذلك بوضوح في نموذج المسـار الـذي صـممه الباحـث، فـالمتغير الـذي يساهم في التأثير لا بد أن يسبق المتغير الذي يقع عليه التأثير، وفي ذلك يستند الباحثون على النظريـات والبحوث السابقة، وكذلك على الرؤية النافذة والحجة والمنطق والفكر وغير ذلك مـن المصـادر؛ بحيـث يعكس نموذج الانحدار تتابعًا زمنيًا مقنعًا للمتغيرات المختلفة، الداخلية والخارجيـة:

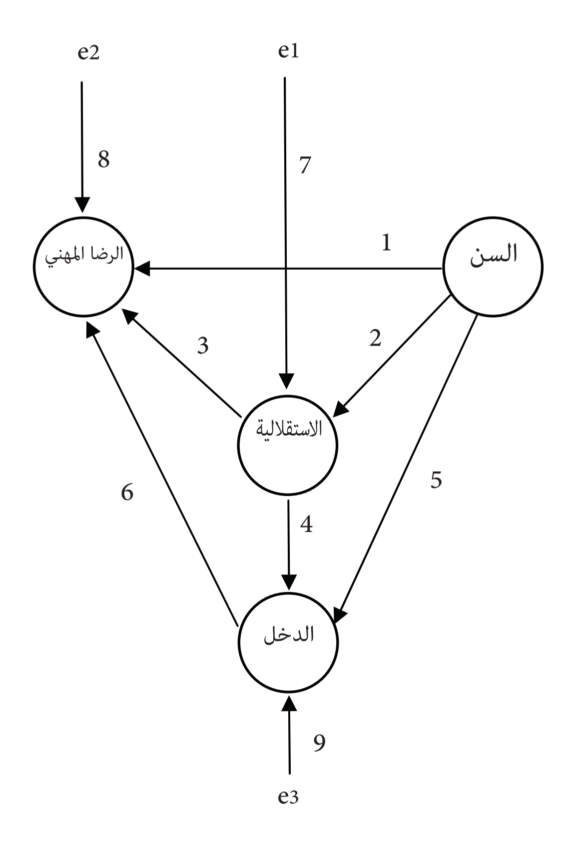


ففي هذا الشكل هناك أربعة متغـيرات (م1، م2، م3، م4)، وقـد تـم ترتيبهـا في تسلسـل زمنـي مـن اليمين إلى اليسار، ويتضمن الشكل أيضًا المتغيرين (ب1)، (ب2)، وهـما متغـيرات البـواقي، ولا يرتبطـان ببعضـهما، ولن يستخدما في التحليل (وذلك بهدف التبسيط).أمـا الخطـوط المسـتقيمة فهـي المسـارات (Paths) التي تمثل التأثيرات السببية (Casual Effects) لكل متغير على الآخر، فالمتغير المـؤثر هـو المتغـير المستقل، أما المتغير الواقع عليه التأثير، فهو المتغير التابع، ومـن الرسـم نتبـين أن المتغـيرين (م1)، (م2) هما متغيران مستقلان يـؤثران تـأثيرًا مبـاشرًا في المتغـير الثالـث (م3)؛ أي أن المتغـير الثالـث (م3) هـو المتغير التابع في علاقته بالمتغيرين الأول والثاني. لكن المتغير الثالث (م3) هو في الوقت نفسـه متغـير مستقل بالنسبة للمتغير الرابع (م4)؛ إذ إن المتغير الثالث يؤثر على المتغير الرابع تأثيرًا مباشرًا.

خطوات تحليل المسار

يبدأ تحليل المسار بتكوين النموذج السببي المفترض (Casual Model)، بمعنى تحديد المتغيرات التي تؤثر في بعضها البعض، وذلك بناء على الدراسات العلمية والنظريات والحقائق، فقد يريد الباحث اختبار ما إذا كان التعرض لوسائل الإعلام يؤثر في المعرفة السياسية، وأن هذه المعرفة تتأثر أيضًا بمتغيرات أخرى (كالمستوى التعليمي والمشاركة السياسية)، وأن متغير الدخل يؤثر في التعرض لوسائل الإعلام.. إن بناء أو تكوين النموذج السببي يعني وضع هذه المتغيرات في علاقة تأثير وتأثر، بناء على اعتبارات منطقية، إنه نموذج تخطيطي ذو اتجاه واحد (Recursive Model) يمثل العلاقة السببية بين المتغيرات، وعلى الباحث أن يحدد طبيعة تأثير كل منها على الآخر سواء كان هذا التأثير مباشرًا أو غير مباشر.

يقصد بالتأثير المباشر أن أحد المتغيرات يؤثر في متغير آخر دون أن يتدخل متغير ثالث في هذا التأثير(كأن نفترض أن التعرض الكثيف لنشرات أخبار التليفزيون يؤثر في المعرفة بالقضايا السياسية)، أما التأثير غير المباشر فيعني أن أحد المتغيرات يؤثر في متغير آخر من خلال متغير ثالث (كأن نفترض أن التعرض الكثيف لنشرات أخبار التليفزيون يؤثر في التذكر، وأن التذكر يؤثر في المعرفة بالقضايا السياسية)، وعلى الرغم من تعدد الكتابات المعنية بنماذج وخطوات تحليل المسار على نحو تفصيلي، إلا أن أهم هذه الكتابات تتمثل فيما قدمه كل بريمان وكرامر (Bryman & Cramer,1999) عن هذا الموضوع، حيث قدما شرحًا توضيحيًا للنموذج السببي المفترض، على النحو التالي:



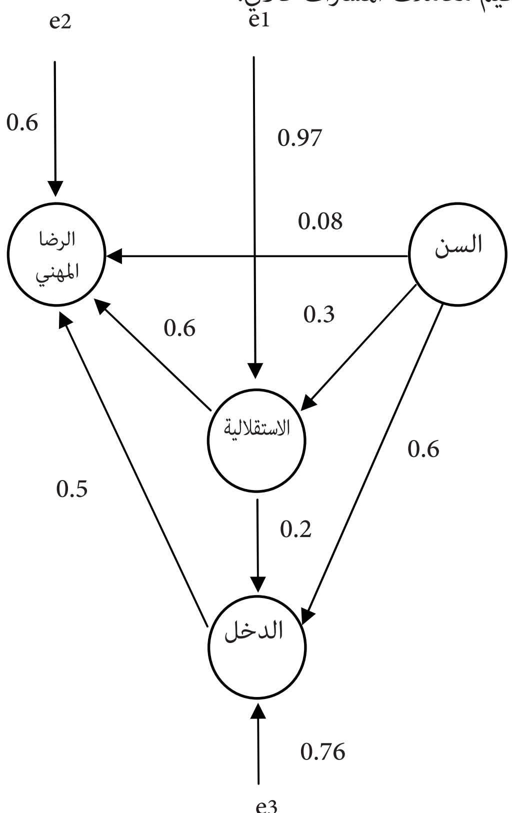
في هذا النموذج لدينا أربعة متغيرات: السن، الدخل، الاستقلالية، الرضا المهني. وتدل الأسهم (Arrows) على العلاقة السببية المتوقعة بين المتغيرات، ويتحرك اتجاهها من اليمين إلى اليسار ليدل على العلاقة السببية التي تتخذ تسعة مسارات يتعين على الباحث أن يجد قيمة كل منها. وحسب هذا النموذج فإن السن له تأثير مباشر في الرضا المهني (المسار الأول)، كما أن السن له تأثير مباشر في الدخل (المسار الخامس)، والدخل يؤثر بدوره في الرضا المهني (المسار السادس)، ويؤثر السن أيضًا في الاستقلالية (المسار الثاني)، كما أن الاستقلالية تؤثر بدورها في الرضا المهني (المسار الثالث). وإذا كان السن يؤثر في الاستقلالية (المسار الثاني)، فإن الاستقلالية تؤثر في الدخل (المسار الرابع)، والدخل بدوره يؤثر في الرضا المهني (المسار الرابع)، كما أن الاستقلالية لها تأثير مباشر في الرضا

المهني (المسار الثالث)، وللاستقلالية أيضًا تأثير غير مباشر في الرضا المهني من خلال تأثيرها المباشر في الدخل (المسار الرابع)، والدخل بدوره ذو تأثير مباشر في الرضا المهني (المسار السادس).

يتضح من النموذج أيضًا أن هناك عوامل خارجية تؤثر في كل من الدخل والاستقلالية والرضا المهني، وهذه العوامل الخارجية تعني التباين غير المفسر (Unexplained Variance)، فالسهم المتجه من (e1) إلى الاستقلالية (رقم 7) يدل على مقدار التباين (Variance) في متغير الاستقلالية وهو التباين الذي لا يمكن تفسيره بمتغير الدخل. أما السهم المتجه من (e2) إلى الرضا المهني (رقم 8) فيدل على مقدار الخطأ (Error) في التباين الخاص بالرضا المهني، هذا الخطأ الذي لا يمكن تفسيره بمتغيرات الدخل والاستقلالية والسن. أخيرًا، فإن السهم المتجه من الخارجي (e3) إلى الدخل يرمز إلى التباين الخاص بالدخل، هذا التباين الذي لا يمكن تفسيره بمتغيرات السن والاستقلالية.

هذه المتجهات (Vectors) الثلاثة (e1; e2 ;e3) إنما تشير إلى العوامل الخارجية (الأخرى) ذات التأثير في الرضا المهني والاستقلالية، والتي لم يتضمنها نموذج تحليل المسار، وهذا في الواقع منطق سليم يعكس حقيقة أنه من الصعب حصر جميع العوامل المؤثرة في الظواهر أو الموضوعات التي تبحثها العلوم الإنسانية والاجتماعية، فالباحث قد يتناول مجموعة من المتغيرات ذات التأثير المفترض في ظاهرة معينة، غير أن هذه الظاهرة قد تتأثر بعوامل ومتغيرات أخرى (غير تلك التي تناولها البحث). وبموجب بناء النموذج السببي المفترض و تحديد المسارات، يتم حساب معاملات المسارات (Paths Coefficients).

إن معامل المسار هو معامل الانحدار المقنن أو المعياري(Standardized Regression Coefficient)؛ أي أن الباحث عليه إيجاد معاملات المسارات باستخدام تحليل الانحدار مع إدخال كل المتغيرات المؤثرة (وبالتالي لا يجوز استخدام طريقة Stepwise)، وبالعودة إلى النموذج السابق فإن المعالجة الإحصائية للبيانات كشفت عن أن قيم معاملات المسارات كالآتي:



ومن هذا الرسم نتبين أن متغير السن له تأثير مباشر ضعيف على الرضا المهني (معامل المسار 0.08)، لكن السن في الوقت نفسه له تأثير قوي غير مباشر على الرضا المهني، فمعامل المسار المتجه من السن إلى الاستقلالية قيمته 0.3 ومعامل المسار المتجه من

الاستقلالية إلى الرضا المهني قيمته 0.6، أما معامل مسار الدخل إلى الرضا المهني فقيمته 0.5وهكذا يمكن توضيح كافة المسارات المبينة في النموذج، فإذا أخذنا مسار السن على الرضا المهني كمثال، نجد أن التأثير المباشر (مسار السن إلى الرضا المهني) قيمته 0.08 أما التأثير غير المباشر، فيتم حسابه من خلال تتبع كل المسارات التي تعكس تأثير السن على متغيرات أخرى تؤثر بدورها في الرضا المهني،فهناك:

- مسار السن، الدخل،الرضا المهني= (0.5) (0.6)= 30%
- مسار السن، الاستقلالية، الرضا المهني = (0.3) (0.6)= 18%
- مسار السن، الاستقلالية، الدخل،الرضا المهني = (0.3) (0.2) (0.5) = 3%

ويكون إجمالي تأثير هذه المسارات الثلاثة = 0.30 + 0.18+ 0.03 = 0.51

هكذا نتبين أنه إذا كان التأثير المباشر لمتغير السن على الرضا المهني قيمته 0.08 فإن التأثير غير المباشر قيمته 0.51 وبالمنطق نفسه يمكن حساب التأثير المباشر والتأثير غير المباشر لكلٍّ من الاستقلالية والدخل على الرضا المهني، ثم المقارنة بين المتغيرات الثلاثة (السن، الاستقلالية، الدخل) من حيث قيمة تأثير كل متغير.

ويفترض تحليل المسار بعض الأسس الإحصائية أهمها ثلاثة أسس: الأساس الأول يتمثل في أن تكون العلاقة بين المتغيرات خطية (Linear) (وبالتالي لا بد من التحقق من شكل العلاقة بين كل متغيرين في النموذج)، أما الأساس الثاني، فهو ألا تختلف العلاقة بين أي متغيرين تبعًا لوجود متغير ثالث (بمعنى ألا يوجد تفاعل بين المتغيرات)، فإذا كانت المتغيرات التي نقوم بدراستها هي: (قراءة الصحف اليومية بانتظام، سماع الراديو يوميّا، مشاهدة الأخبار في التليفزيون يوميّا)، فإن العلاقة بين أي متغيرين يجب ألّا تزيد أو تقل بتأثير المتغير الثالث. ويمكن التأكد من ذلك إما من خلال الكشف الآلي عن التفاعل بين المتغيرات باستخدام (Automatic Interaction Detection) (AID) الذي تتيحه

البرامج الإحصائية الجاهزة (كما يمكن استخدام معامل الارتباط الجزئي). أخيرًا، فإن الأساس الثالث الذي يقوم عليه تحليل المسار هو ألا تـرتبط متغـيرات البـواقي (Residuals) ببعضـها الـبعض أو بغيرهـا مـن المتغيرات في نموذج تحليل المسار الذي صممه الباحث.

إن أي نموذج سببي (Casual) يشتمل عـلى بعـض الأخطـاء أو البـواقي؛ ونظـرًا لأن تحليـل المسـار يعتمد على تحليل الانحدار المتعدد، فإنه يفترض فيه أن معاملات الارتباط بين البواقي وجميع المتغيرات الخارجية تساوي الصفر. وتأتي المشكلة من أن الترتيب الزمني للمتغـيرات كـما يعـبر عنـه النمـوذج قـد يكون معيبًا أو غير دقيق، وبالتالي يكون نموذج تحليل المسار الذي تم تصميمه معيبًا، بل ومضللًا (لأنـه على سبيل المثال إذا تم ترتيب المتغيرات ترتيبًا غير صحيح من الوجهة السببية سيعني ذلك أن تحديد المتغيرات الداخلية والخارجية غير صحيح، ويترتب على ذلك خطأ تقـدير معـاملات المسـارات وخطـأ في تقدير بواقي المتغيرات التي لم توضع في ترتيبهـا الصـحيح). وإذا كـان تحليـل المسـار يتـيح الكثـير مـن النتائج، إلا أن ذلك مقيد بدقة النموذج النظري للمسارات، وهذا النموذج رغم اعتماده عـلى البحـوث والنظريات العلمية، إلا أنه قد لا يكون دقيقًا، ولا بد من أخذ هذه الحقيقة في الاعتبـار عنـد اسـتخدام تحليل المسار.

المبحث السادس

التحليل التمييزي

على الرغم من أن تحليل التباين المتعدد (MANOVA) يمكن من فحص الفروق بين المجموعات في عـدة متغـيرات، إلا أنـه لا يبـين مـا إذا كانـت هـذه الفـروق ذات أهميـة في التمييـز الحقيقـي بـين المجموعات المختلفة، ومن هنا يستخدم البـاحثون التحليـل التمييـزي (Discriminant Analysis) الـذي يتيح تكوين متغيرات جديدة من المتغيرات الأصلية (التي تتمثل فيها الفـروق بـين المجموعـات)، هـذه المتغيرات الجديدة تسمى الدوال التمييزية (Discriminant Functions).

فمن المعروف إحصائيًا أن الدالة (Function) هي صيغة رياضية تحدد احتمال أن يتخذ المتغير العشوائي قيمة معينة، والمتغير العشوائي هو أي كمية (Quantity) تتحدد قيمتها بالصدفة. وفي التحليل التمييزي تتكون الدالة من تجميعات خطية(Linear Combinations) للمتغيرات المنبئة (Predictor Variables) التي توضح التمييز الدقيق بين المجموعات. فالتحليل التمييزي يتم من خلاله بناء أو تكوين نموذج تنبؤي بانتماء المبحوث إلى مجموعة معينة حسب استجاباته وخصائصه من حيث المتغيرات المنبئة.وقد يكون التصنيف على أساس مجموعتين فقط، كما قد يكون على أساس أكثر من مجموعتين (عندما يكون التصنيف إلى مجموعتين فقط، فإن التحليل التمييزي يسفر عن دالة واحدة، أما عندما يكون التصنيف إلى أكثر من مجموعتين فإن التحليل التمييزي يسفر عن أكثر من دالة).

وكمثال توضيحي، نفرض أن أحد الباحثين كشفت دراسته عن أن «العاطلين عن العمل» يشاهدون التليفزيون بدرجة أكبر مقارنة «بالعاملين»، وأن العاملين يستخدمون الإنترنت بدرجة أكبر مقارنة بالعاطلين. غير أن هذا لا يكفي «للتمييز» بين مجموعة «العاطلين» ومجموعة «العاملين»؛ لأن هناك متغيرات أخرى لا بد من أخذها في الاعتبار إذا أردنا تمييزًا دقيقًا بين المجموعتين، من هذه المتغيرات على سبيل المثال السن ومستوى التعليم.

في هذه الحالة، فإن الدالة التمييزية تشمل متغيرات: مشاهدة التليفزيون، استخدام الإنترنت، السن، مستوى التعليم. فإذا كانت هذه المتغيرات ذات تأثير جوهري في التمييز بين المجموعتين، فإن قيمة هذه الدالة سوف تختلف اختلافًا جوهريًا بين المجموعتين والعكس صحيح (بمعنى أن هذه المتغيرات إذا لم تكن ذات تأثير جوهري في التمييز بين المجموعتين، فإن قيمة الدالة لن تختلف اختلافًا جوهريًا بين المجموعتين).

ويتم تقـدير معاملات الدالة التمـييزية الخطية كمعادلة الانحـدار الخطـي المتعـدد Multiple) (A) Linear Regression Equation) فإذا رمزنا إلى قيم هذه المعاملات بالرموز: س، ص، ع، ل فإن قيمة الدالة التمييزية تساوي:

(س × مشاهدة التليفزيون) + (ص× استخدام الإنترنت)

+ (ع × السن) + (ل× التعليم)

مرة أخـرى، إذا اختلفـت قيمـة هـذه الدالـة في حالـة مجموعـة «العـاطلين» عـن قيمتهـا في حالـة مجموعة «العاملين»، يمكن القول بأن المتغيرات المذكورة «تميز» بين المجموعتين.

وتتضـح أهميـة كـل دالـة تمييزيـة **عـلى** حـدة مـن واقـع قيمـة الارتبـاط الحقيقـي Canonical) Correlation) (والذي يسمى أحيانًا بالارتباط القانوني أو الارتباط المعترف به) ومن النسبة المئوية للقوة التمييزية التي تحققها الدالة. ففيما يخص «الارتباط الحقيقي»، فإنه يعبر عن قوة العلاقـة بـين الدالـة التمييزية وبين الانتماء للمجموعات، فإذا كان الارتباط الحقيقـي للدالـة الأولى مـثلًا 0.68 وكـان للدالـة الثانية 0.53، فإن هذا يعني أن الدالة الأولى تميز بين المجموعات أكبر من الدالة الأولى، أما فـيما يخـص النسبة المئوية للقوة التمييزية فإنها تعبر عن نسبة التباين بين متوسطات المجموعـات الراجعـة للدالـة إلى نسبة التباين الذي توفره كل الدوال، فإذا كانت نسبة القوة التمييزيـة للدالـة الأولى تسـاوي 75% مثلًا، فإن هذا يعني أن هذه الدالة تفسر 75% من التباين الإجمالي الذي تفسـره كل الدوال التمييزية التي استخرجها النموذج، وعلى هذا الأساس يمكن معرفة القوة التمييزية لكل دالة.

كما أن معاملات الارتباط الحقيقي (Canonical Correlation) هي في جوهرها توضـيح للارتباط بين متغيرات الدراسة والدوال التمييزيـة. نفـترض مـثلًا أن المعطيـات الإحصـائية تفيد بأن معامل الارتبـاط الحقيقـي بـين الدالـة الأولى ومتغـير «المعرفـة السياسـية» هـو 0.78، وأن الارتباط بين الدالـة الأولى ومتغـير «الاهـتمام بالموضـوعات السياسـية» هـو

0.65 كما أن الارتباط بين الدالة الأولى ومتغير «حضور اللقاءات السياسية» هو 0.35، هذا يعني أن الدرجة الأكبر على الدالة الأولى تعكس المعرفة السياسية، كما يمكن تسمية هذه الدالة (المشاركة السياسية). وتتصف المعطيات الإحصائية للتحليل التمييزي بالثراء الشديد؛ الأمر الذي يتطلب من الباحث التعامل معها بوعي واستخدامها في الدراسة الاستخدام الصحيح،ومن أبرز هذه المعطيات:

- إحصاءات عن كل متغير، ومنها:

Means, Standard Deviations, Univariate ANOVA.

- مخرجات إحصائية عن كل خطوة (إجراء)، مثل:

Box's M, Within-Groups Correlation Matrix, Within-Groups Covariance Matrix, Separate-Groups Covariance Matrix, Total Covariance Matrix.

- إحصاءات عن كل دالة تمييزية، ومنها :

Eigenvalue Percentage of Variance, Canonical Correlation, Wilks' Lambda, Chi-Square. For Each Step: Prior Probabilities, Fisher's Function Coefficients, Un-Standardized Function Coefficients, Wilks' Lambda for Each Canonical Functio.

كما تتيح المعطيات الإحصائية التي يسفر عنها التحليل التمييزي مفهوم (Group Centraids)، ويعني مراكز المجموعات على الدوال التمييزية، وهذه المراكز هي المتوسطات الحسابية لدرجات المجموعات على الدوال التمييزية بعد تحويلها إلى درجات معيارية (Z Scores)، فإذا قلنا بأن درجات المجموعة الأولى على الدالة التمييزية الأولى هو 0.712 بينما درجة المجموعة الثانية على الدالة نفسها هو 1.56، فإن ذلك يعني أن المجموعة الثانية تتفوق على المجموعة الأولى بمقدار 0.848؛ أي أقل من واحد انحراف معياري.

ومـن الضروري عنـد اسـتخدام التحليـل التمييـزي أن يـتم تكويـد المجموعـات بأرقـام صـحيحة (Integers)، وعند إجراء التحليل يتم تحديد الحـد الأدنى والحـد الأقصـى لهـذه الأرقـام؛ إذ إن التحليـل يستبعد تلقائيًا؛ أي حالة (مبحوث) خارج هذه الحدود، فإذا كان لدينا ثلاث مجموعـات مـثلًا، تكـون أرقامها الكودية (1، 2، 3)، فإذا كان هناك أي حالة تم تكويدها برقم آخـر- عـلى سـبيل الخطـأ مـثلًا-، فإنها لن تكون ضمن التحليل؛ الأمر الذي يؤكد ضرورة فحص الاستجابات البسيطة بدقة قبل المعالجـة الإحصائية.

المبحث السابع

التحليل البعدي والتحليل الثانوي

هناك الكثير من البحوث بالغة الأهمية والعمق والثراء الفكري والمنهجي التـي قامـت عـلى إعـادة التحـليل لنتـائج وبيانات إحصائية سابقة، وتستخدم تلك البحوث أسـاليب إحصـائية متطورة، وتقـوم على التحليل البعدي والتحليل الثانوي، ومن الجدير بالتأكيد عليه هـو أن هنـاك معـانيَ كثـيرة لهـذين النمطين من التحليل، وفيما يلي توضيح لكلٍّ منهما حسب الرؤية الأكثر شيوعًا:

أولًا: التحليل البعدي

إن التحليل البعدي (Meta Analysis) في معناه العـام هـو اسـتخدام نتـائج توصـلت إليهـا بحـوث ودراسات سابقة في صياغة رؤية علمية متكاملة عن موضوع معـين؛ أي أن الفكـرة الأساسـية في مـنهج التحليل البعدي هي فكرة تكامـل البحـوث(Research Integration) والاسـتفادة منهـا في إنتـاج دراسـة جديدة ذات طبيعة كمية (Quantitative) ترتكز على خلفية نظرية وتفسير متعمق.

يعود الفضل في ابتكار التحليل البعدي (Meta Analysis) إلى العالم الأمريكي جلاس (Glass,1976) والذي قال: «إن التحليل البعدي هو (تحليل التحليل)»، بمعنى التحليل الإحصائي لمجموعة كبيرة من النتائج التي توصلت إليها دراسات سابقة، وبموجب هذا التحليل نحصل على معطيات جديدة؛ ومن ثمّ صياغة هذه المعطيات (الجديدة) في تقرير علمي متكامل. ويتطلب التحليل البعدي التسجيل الكمي لخصائص الدراسات السابقة ونتائجها والتعامل مع المادة المسجلة على أنها بيانات (Data) تخضع لمعالجة إحصائية جديدة يمكن الحصول منها على نتائج جديدة/متكاملة بما يقابل أهداف البحث المطلوب إنتاجه من الدراسات السابقة. مثال ذلك أن يقوم الباحث برصد نتائج بحوث تحليل المحتوى التي أجريت على الأخبار في التليفزيون المصري خلال مدة زمنية معينة؛ ومن ثمّ يقوم برصد نتائج تلك البحوث وإدخالها في الحاسوب ومعالجتها إحصائيًا بطريقة تتيح التوصل إلى نتائج جديدة تتفق وهدف الدراسة التي يقوم بها، وقد يقوم الباحث برصد نتائج البحوث التي تضمنت تقصي العلاقة بين النوع ومشاهدة التليفزيون خلال عشر سنوات مثلًا، ومعالجة تلك النتائج والخروج من ذلك بتقرير متكامل عن الفروق بين الجنسين في مشاهدة التليفزيون.

ويمكن تنفيذ التحليل البعدي بكفاءة عالية إذا توفرت لدى الباحث أهداف واضحة وقابلة للتنفيذ، ثم القدرة على تطويع نتائج الدراسات السابقة بما يتفق مع تلك الأهداف، ويتطلب ذلك التكويد الكمي (Quantitative Coding) الصحيح لنتائج الدراسات السابقة والعينات التي أجريت عليها، هذا بالإضافة إلى خصائص تلك الدراسات؛ من حيث تاريخ النشر ومصدره وخصائص المفحوصين وطبيعة المعالجة وتصميم الدراسات وطرق القياس، وغير ذلك من الخصائص التي يراها الباحث ضرورية للبحث المزمع القيام به، وفي بعض بحوث التحليل البعدي يكون من الضروري تحويل نتائج الدراسات السابقة إلى نظام قياسي مشترك حتى يمكن التعامل معها إحصائيًا والتوصل إلى نتائج دقيقة غير مضللة.

ثانيًا: التحليل الثانوي

التحليل الثانوي (Secondary Analysis) هو الاستعانة بالبيانات الخام (Raw Data) الخاصة بدراسات سابقة، وإعادة تحليل هذه البيانات إحصائيًا من منظور معين، إما للإجابة على تساؤلات، أو للتحقق من فروض معينة (على أساس أن تلك البيانات تتضمن كافة المتغيرات اللازمة)، أو للحصول على نتائج يمكن مقارنتها؛ أي أنه في التحليل الثانوي لا يتم جمع البيانات من مصادرها الأولية المباشرة، ولكن يتم الحصول على البيانات الخاصة بدراسات سابقة، وقد تكون هذه البيانات مخزونة في الحاسوب؛ وفي حالة فقدان هذه البيانات المخزونة لسبب أو لآخر يمكن الرجوع إلى الاستبيانات الأصلية في حالة وجودها. وقد انتشر استخدام التحليل الثانوي مع تطور الطرق الإحصائية واحتياجات البحوث، وتراكم البيانات الخام من الدراسات السابقة.

ومن خلال إعادة المعالجة الإحصائية للبيانات الخام (Raw Data) الخاصة بدراسات سابقة، يمكن التوصل إلى نتائج جديدة، وكثيرًا ما يفيد هذا المدخل في بلورة رؤية جديدة، كما يفيد في التأكد من نظرية، أو التحقق من عدة فروض إثباتًا أو نفيًا، وكذلك يفيد في توضيح الفروق أو العلاقات بين متغيرات معينة، أو رصد ملامح التغير على موضوع ما عبر الفترات الزمنية المختلفة، فالباحث قد يحصل على البيانات الخام الخاصة بمشاهدة الإعلانات التليفزيونية من واقع دراسات أجريت على امتداد عشرين عامًا، ثم يعالج هذه البيانات إحصائيّا؛ بهدف التوصل إلى نتائج توضح تطور العلاقة بين مشاهدة إعلانات التليفزيون والسلوك الشرائي للجمهور.

خلاصة الفصل الخامس

تضمن هذا الفصل تعريفًا بالجوانب الأساسية لبعض أساليب التحليل الإحصائي المتعمقة، وتتلخص موضوعاته الأساسية في:

- يستخدم تحليل التباين المتعدد (MANOVA) لتقصي الاختلافات أو الفروق بين المجموعات في أكثر من متغير تابع، وهناك عدة شروط يتعين توافرها حتى يتسنى استخدام (MANOVA) لا بد أن يعيها الباحث جيدا.
- من خلال التباين المتعدد يمكن معرفة نسبة التباين في المتغيرات التابعة التي يمكن تفسيرها بالمتغير المستقل، ويعبر عن ذلك في المعطيات الإحصائية بقيمة تربيع إيتا الجزئي (Partial ETA. Square)، فإذا كانت هذه القيمة مثلًا 0.22، فإن ذلك يعني أن المتغير المستقل يفسر 22% من التباين في المتغير التابع.
- تحليل التغاير (ANCOVA) (Covariance Analysis) هو أسلوب إحصائي للتحكم في المتغيرات التي تتداخل مع المتغير المستقل في التأثير على المتغير التابع.هذه المتغيرات التي يتم ضبطها أو التحكم فيها تسمى بالمتغيرات المصاحبة أو الملازمة (Covariates)، ويعتبر تحليل التغاير من أساليب الضبط الإحصائي الهامة، كما أنه يستخدم أيضًا بهدف زيادة قوة اختبار تحليل التباين الأحادي (ANOVA)، وفهم التباين غير المتنبأ به أو غير المتوقع في نموذج تحليل التباين، وهو تباين الخطأ (التباين داخل المجموعات)؛ الأمر الذي يؤدي إلى نتائج أكثر دقة.
- تحليل الانحدار (Regression Analysis) هو أسلوب إحصائي يمكن بواسطته التنبؤ بقيمة متغير معين بمعلومية قيمة متغير آخر بناءً على وجود ارتباط جوهري بين المتغيرين. وخط الانحدار هو ذلك الخط الذي يجعل مجموع الانحرافات عنه

أقل ما يمكن، سواء فيما يخص المتغير (س) أو فيما يخص المتغير (ص)، ويمكن من خلاله تحديد أي نقطة على المحور الصادي إذا علمنا النقطة المقابلة لها على المحور السيني، والعكس صحيح.

- الانحدار الخطي البسيط هو التطبيق البسيط أو المباشر للفكرة الأساسية التي يقوم عليها الانحدار؛ فهو يهدف إلى التنبؤ بقيمة متغير معين بمعلومية متغير آخر؛ أي أنه يتضمن متغيرين فقط أحدهما هو المتغير المستقل والثاني هو المتغير التابع.

- انحدار (ص) على (س) يعني التنبؤ بقيمة المتغير (ص) بناء على معرفة قيمة (س)، أما انحدار (س) على (ص) فيعني التنبؤ بقيمة المتغير (س) بناء على معرفة قيمة (ص).

- إن الانحدار المتعدد (Multi regression) يتعامل مع أكثر من متغيرين، فقد يكون لدينا أكثر من متغير مستقل، متغيران مستقلان أو ثلاثة متغيرات مستقلة أو أكثر، ونريد معرفة تأثيرها في متغير تابع، هنا نستخدم الانحدار المتعدد؛ حيث يكون التعامل مع أكثر من متغير مستقل (More Than Independent Variable).

- معامل الانحدار المتعدد يوضح لنا مقدار التغير الذي يساهم به كل متغير مستقل على حدة في المتغير التابع، كما يكشف لنا عن مجمل التغير الذي تحدثه جميع المتغيرات المستقلة (مجتمعة) في المتغير التابع.

- قبل تطبيق الانحدار المتعدد يجب التأكد أولًا من أن الارتباط بين كل متغيرين من المتغيرات المستقلة ليس قويًا (فلا يتجاوز 0.8)، علمًا بأن قيمة معامل الانحدار الخاص بمتغير مستقل معين لا يدل في حد ذاته على أهمية هذا المتغير في المتغير التابع، وذلك إذا كانت وحدات كل متغير مختلفة عن وحدات المتغير الآخر، ومن هنا، يتم

الحصـول عـلى وزن بيتـا (Beta weight) أو معامـل الانحـدار المقـنن (Standardized Beta Coefficient).

- إن الانحدار اللوجستي (Logistic Regression) يبحـث في معـدل التغـير في لوغـاريتم الأرجحيـة (Log Odds) لحدوث المتغير التابع Y عندما يتغير المتغـير المسـتقل X. إن نسـبة الأرجحيـة هـي احتمال وقوع الحدث مقسومًا على احتمال عدم وقوعه.

- يستخدم الانحدار اللوجستي (Logistic Regression) عادة عندما يكون المتغير التابع ثنائيّـا متقطعًـا (Binary Dichotomous)، بمعنى أن هذا المتغير له حالتان فقط، ويمكن إعادة تكويد المتغير المسـتمر (غير المتقطع) بحيث يصبح متغيراً ثنائيّا متقطعًا.

- التحليل العاملي (Factor Analysis) مـن الأسـاليب الإحصـائية المهمـة التـي يمكـن أن يسـتخدمها الباحث في تحليل البيانات متعـددة المتغـيرات، ودراسـة العلاقـات القائمـة بـين تلـك المتغـيرات واختزالها في عدد أقل من العوامل(Factors) التي تفسـر الظاهرة مجال الدراسة.

- إن الدراسة العاملية لا بد أن ترتكز على أهداف وفـروض / تسـاؤلات واضـحة ومحـددة وقابلـة للقياس وذات دلالة لموضوع الدراسـة وأهـدافها وإطارهـا النظـري، كـما أن الدراسـة العامليـة تتطلب التحديد الدقيق للمتغيرات وكفاءة اختيار مفردات العينة والحجم المناسب، وكذلك دقة تصميم أداة / أدوات القياس.

- هناك مجموعة من المفاهيم الأساسية الخاصة بالتحليـل العـاملي، ولا بـد أن يسـتوعبها الباحـث، وهذه المفاهيم هي :

- العامـــل (Factor)، وهـــو المعنـــى العـــام الـــذي يعـــبر عـــن البنـــود ذات الارتبـــاط بعضها ببعض.

- مصفوفة الارتباط (Correlation Matrix) وهي قيم معاملات الارتباط بين المتغيرات بعضها بعضًا، أو العوامل بعضها بعضًا، أو بين العوامل والمتغيرات.

- الاشتراكيات أو قيم الشيوع وهي التباين المشترك (Common Variance).

- التشبعات (Saturations)، وهي ارتباط المتغير بالعامل (Factor)، الجذر الكامن (Eigenvalue أو Latent Root)، وهو حجم التباين الذي يفسره العامل، وهو يدل على مجموع مربعات تشبعات هذا العامل بالبنود التي يتشبع بها.

- هناك نمطان رئيسيان للتحليل العاملي، النمط الأول هو التحليل العاملي الاستكشافي (Exploratory Factor Analysis)، والنمط الثاني هو التحليل العاملي التأكيدي (Confirmatory Factor Analysis)، وتتعدد طرق التحليل العاملي، كما أنها تتطور باستمرار، ولكلٍّ من هذه الطرق أسسها الرياضية.

- يقصد بتدوير العوامل أن يحل العامل محل المتغيرات (البنود)؛ بحيث تصبح التشبعات بدائل للدرجات المعيارية، أما عن عملية التدوير ذاتها، فتعني تدوير العوامل حول نقطة الأصل حتى تصل إلى وضع بديل.

- بعد استخلاص العوامل ينبغي تفسير البنية العاملية الناتجة عن التحليل، وإضفاء معنى على العوامل بحسب مجال الدراسة، وذلك بتسمية العوامل تسمية مناسبة، والاستعانة بالتمثيل البياني لتوضيحها واستيعابها، وتعتمد تسمية العوامل (Naming Factors) على الهدف من التحليل ومنظور الباحث.

- تحليل المسار (Path Analysis) هو أسلوب في تحليل البيانات لاختبار صحة نموذج يفترضه الباحث لتفسير العلاقات بين المتغيرات موضع البحث بطريقة متسقة، فإذا افترض الباحث (بناء على أدلة علمية) أن المتغير (س) يؤثر في المتغير (ص) الذي

يؤثر بدوره في المتغير (ع)، فإن استخدام تحليل المسار يمكن أن يثبت صحة هذا النموذج أو خطأه. أي أن تحليل المسار ليس طريقة للكشف عن العلية أو السببية (Causality)، وإنما هو طريقة لاختبار نموذج سببي (Casual Model) يفترضه الباحث، إنه يقول لنا: «تأثير عدة متغيرات على بعضها بعضًا»، ومن خلال تحليل المسار يمكن تقدير التأثير النسبي للمتغيرات ضمن هذا النموذج المفترض.

- يبدأ تحليل المسار بتكوين النموذج السببي المفترض (Casual Model)، بمعنى تحديد المتغيرات التي تؤثر في بعضها بعضًا، وذلك في ضوء الدراسات العلمية والنظريات والحقائق. وفي بناء نموذج تحليل المسار لا بد من مراعاة حقيقة أن السبب لا بد أن يسبق النتيجة أو الأثر (The Cause Must Precede the Effect)، وبالتالي لا بد من مراعاة الترتيب الزمني للمتغيرات التي تحت الدراسة؛ بحيث يظهر ذلك بوضوح في نموذج المسار الذي صممه الباحث.

- يستخدم التحليل التمييزي (Discriminant Analysis)؛ بهدف تكوين متغيرات جديدة من المتغيرات الأصلية (التي تتمثل فيها الفروق بين المجموعات)، هذه المتغيرات الجديدة تسمى الدوال التمييزية (Discriminant Functions). والدالة (Function) هي صيغة رياضية تحدد احتمال أن يتخذ المتغير العشوائي قيمة معينة، والمتغير العشوائي هو أي كمية (Quantity) تتحدد قيمتها بالصدفة. وفي التحليل التمييزي تتكون الدالة من تجميعات خطية (Linear Combinations) للمتغيرات المنبئة (Predictor Variables)، والتي توضح التمييز الدقيق بين المجموعات. فالتحليل التمييزي يتم من خلاله بناء أو تكوين نموذج تنبؤي بانتماء المبحوث إلى مجموعة معينة حسب استجاباته وخصائصه من حيث المتغيرات المنبئة.

- إن التحليل البعدي (Meta Analysis) في معناه العام هو استخدام نتائج توصلت إليها بحوث ودراسات سابقة في صياغة رؤية علمية متكاملة عن موضوع معين؛ أي أن الفكرة الأساسية في منهج التحليل البعدي هي فكرة تكامل البحوث (Research Integration)، والاستفادة منها في إنتاج دراسة جديدة ذات طبيعة كمية (Quantitative) ترتكز على خلفية نظرية وتفسير متعمق.

- التحليل الثانوي (Secondary Analysis) هو الاستعانة بالبيانات الخام (Raw Data) الخاصة بدراسات سابقة، وإعادة تحليل هذه البيانات إحصائيًا من منظور معين، إما للإجابة على تساؤلات، أو للتحقق من فروض معينة (على أساس أن تلك البيانات تتضمن كافة المتغيرات اللازمة)؛ أو للحصول على نتائج يمكن مقارنتها.

مصادر الفصل الخامس ومراجعه

(أ) مصادر ومراجع عربية:

- رجاء محمود أبو علام (1999)، مناهج البحث في العلوم النفسية والتربوية، (القاهرة: دار النشر للجامعات).
- زكريا الشربيني (1990)، الإحصاء اللابارامتري في العلوم النفسية والاجتماعية والتربوية، (القاهرة: مكتبة الأنجلو المصرية).
- سعدي شاكر حمودي (2000)، علم الإحصاء وتطبيقاته في المجالين التربوي والاجتماعي، (عمان: مكتبة دار الثقافة للنشر والتوزيع).
- سمير كامل عاشور& سامية أبو الفتوح سالم (2003)، العرض والتحليل الإحصائي باستخدام (SPSSWIN)، الجزء الأول، (القاهرة: جامعة القاهرة: معهد البحوث والدراسات الإحصائية).
- سمير كامل عاشور& سامية أبو الفتوح سالم (2005)، العرض والتحليل الإحصائي باستخدام (SPSSWIN)، الجزء الثاني، (القاهرة: جامعة القاهرة: معهد البحوث والدراسات الإحصائية).
- صلاح الدين محمود علام (2000)، تحليل بيانات البحوث النفسية والاجتماعية والتربوية، (القاهرة: دار الفكر العربي).
- عاطف أحمد منصور(1992)، الرياضيات المسلية، (القاهرة: مكتبة ابن سينا).
- عبد الحميد محمد نجم & محمد عبد الهادي المحميد (1990)، الإحصاء الوصفي والتحليلي مع استخدام البرامج الجاهزة، (الكويت: مكتبة جامعة الكويت).

- عزت عبد الحميد محمد حسن (2011)، الإحصاء النفسي والتربوي، (القاهرة: دار الفكر العربي).

- فتحي عبد الله فياض (1991)، التحليل الإحصائي للبيانات الجغرافيـة، (القـاهرة: مكتبـة جامعـة عين شمس).

- فؤاد أبو حطب & آمـال صـادق (2010)، منـاهج البحـث وطـرق التحليـل الإحصـائي في العلـوم النفسية والاجتماعية والتربوية، (القاهرة: مكتبة الأنجلو المصرية).

- فؤاد البهي السيد (1978)، علم النفس الإحصائي، (القاهرة: دار الفكر العـربي).

(ب) مصادر ومراجع أجنبية:

- Ahuvia, Aaron (2001), Traditional, Interpretive, and Reception Based Content Analyses: Improving the Ability of Content Analysis to Address Issues of Pragmatic and Theoretical Concern. Social Indicators Research, VOL. 54 (2), pp. 139 -172.

- Barrow Michael (2001), Statistics for Economics, Accounting and Business Studies. London. Longman.

- Bars, Donald (2004), The Mixed General Linear Model of Statistical Analysis and EEG. Journal of Nero Therapy, Vol. 8 (2), 2004. pp.132-133.

- Bekoff,-Marc; Allen, - Colin; Grant,- Michael - C. (1999), Feeding Decisions by Steller's Jays (Cyanocitta Stelleri): The Utility of a Logistic Regression Model for Analyses of Where, What and With Whom to Eat. ETHOLOGY. Vol.. 105 (5): 393-406.

- Cameron, Debora (2001), Working With Spoken Discourse. London. Sage Publication.

- Chen,-Wen-Hung; Thissen,-David (1999), Estimation of Item Parameters for the Three-Parameter Logistic Model Using the Marginal Likelihood of Summed Scores. British-Journal-of-Mathematical-and-Statistical-Psychology. Vol. 52 (1): 19-37.

- Cooper, Harris M.; Lindsay, James J. (1998), Research Synthesis and Meta-Analysis.; In: Handbook of Applied Social Research Methods. Bickman, Leonard; Rog, Debra J.; Thousand Oaks, CA, US: Sage Publications, Inc, pp. 315-337.

- Fan,-Xitao; Wang,-Lin (1999), Comparing Linear Discriminant Function with Logistic Regression for the Two-Group Classification Problem. Journal – of - Experimental-Education. Vol. 67(3): 265-286.

- Fife-Schaw, Chris (2000), Introduction to Structural Equation Modeling. In: Research Methods in Psychology (2nd ed.). Breakwell, Glynis M.; Hammond, Sean; Fife-Schaw, Chris; Thousand Oaks, CA: Sage Publications Ltd, pp. 397-413.

- Glass, G. V. (1976), Primary, Secondary and Meta Analysis in Education and Psychology. EDU. Researcher. VOL. 5, pp3-8.

- Robins,-Garry; Pattison, -Philippa; Wasserman, -Stanley (1999), Logit Models and Logistic Regressions for Social Networks: III. Valued Relations. Psychometrika. Vol. 64 (3): 371-394.

- Schmidt, Frank L.; Hunter, John E. (2003), Meta-Analysis. In: Handbook of Psychology: Research Methods in Psychology, Vol.. 2. Schinka, John A.; Velicer, Wayne F.; Hoboken, NJ, US: John Wiley & Sons, Inc, pp. 533-554.

- Schumm, Walter R. (2004), Did the Defeat of Saddam Hussein Reduce Suicide Bombing Casualties and Attacks in Israel? A Statistical Analysis. Psychological Reports, Vol. 95 (3, Part1), pp. 831-834.

- Wampler, Karen S.; Reifman, Alan; Serovich, Julianne M. (2005), Meta-Analysis in Family Therapy Research. In: Research Methods in Family Therapy (2nd ed.). Sprenkle, Douglas H.; Piercy, Fred P.; New York, NY, US: Guilford Press, pp. 318-338.

- Zhu, Hong-Tu; Lee, Sik-Yum (1999). Statistical Analysis of Nonlinear Factor Analysis Models British Journal of Mathematical and Statistical Psychology, Vol. 52 (2), pp. 225-242.

الفصل السادس

العينات في البحث العلمي

تمهـــيد

من المعروف أن العينة (Sample) جانب جوهري في البحث العلمي بكافة فروعه وتطبيقاته، وقد اهتمت أدبيات مناهج البحث وعلم الإحصاء بالعينات من زوايا متعددة، بل إن العينات تخصص عميق ضمن هذا العلم . ونظرًا لما تشكله العينة في البحوث الاجتماعية من أهمية خاصة، فإن الفصل الحالي يتضمن تعريفًا موجزًا بالعينات؛ من حيث مفهومها وخصائصها، بما في ذلك حجم العينة وطرق تقديره، والعوامل المؤثرة فيه وكيفية تحديده مع توضيح كيفية استخدام الجداول الإحصائية التي تساعد الباحثين في هذا الشأن، كما يتضمن الفصل تعريفًا بطرق المعاينة العشوائية (العينة العشوائية البسيطة، العينة العشوائية المنتظمة، العينة العشوائية الطبقية، ثم العينة العشوائية متعددة المراحل)، وكذلك طرق المعاينة غير العشوائية (العينة العارضة، والعينة المقصودة، وعينة الحصص). ويختتم الفصل بتوضيح أخطاء المعاينة لما لها من أهمية في الضبط الإحصائي وإجراءات التقدير والاستدلال على خصائص المجتمع بناءً على دراسة العينة.

المبحث الأول

مفهوم وحجم العينة

أولًا: مفهوم العينة

العينة (Sample) هي جزء من المجتمع تتوافر فيه خصائص هذا المجتمع، فإذا أردنا دراسة عينة من خمسة تلاميذ في فصل دراسي يضم أربعين تلميذًا، فإن هؤلاء التلاميذ الخمسة يشكلون العينة ويرمز لها بالحرف (n) (الصغير). أما الأربعون تلميذًا فهم يشكلون المجتمع (Population)؛ أي الأصل الذي سحبت منه العينة، ويرمز له بالحرف (N) (الكبير)، وإذا كانت العينة ينبغي أن تتوافر فيها خصائص المجتمع الذي سحبت منه، فإن الحكمة من إجراء الدراسة على العينة تتمثل في أنه في كثير من الأحيان يستحيل إجراء الدراسة على المجتمع، وبالتالي يتم اختيار عينة محدودة ودراستها بهدف التوصل إلى نتائج يمكن تعميمها على المجتمع. ويصبح ذلك ممكنًا، إذا كانت خصائص العينة تمثل خصائص المجتمع من حيث أكبر عدد ممكن من المتغيرات، خاصة المتغيرات التي يحتمل أن تؤثر في الظاهرة محل البحث.

افترض أن مجتمع البحث يتكون من آلاف من التلاميذ وليكن عشرين ألفًا، في هذه الحالة قد يكون من غير الممكن إجراء الدراسة على هذا العدد، وهنا يتم اختيار عينة من التلاميذ بحيث تتوافر فيها خصائص المجتمع، فإذا كان مجتمع البحث يتوزع حسب متغير الجنس بين الذكور بنسبة 60% والإناث بنسبة 40%، ويريد الباحث أن يكون تركيب العينة متناسبًا مع تركيب المجتمع من حيث توزيع الجنسين، فإن العينة تعكس ذلك بسهولة، فإذا كان حجم العينة هو (100) مفردة مثلًا، فهذا يعني أن العينة سوف تتضمن (60) من الذكور، مقابل (40) من الإناث، مع الاختيار العشوائي للمفردات من

الجنسين .. وهكذا في بقية المتغيرات التي قد تتمثل في الصف الدراسي، ومحل الإقامة، ومستوى دخل الأسرة ... إلخ.

والعينات في البحث الإعلامي - كغيره من البحوث العلمية - يفترض أن تكون ممثلة للمجتمع حتى يمكن تعميم نتائج البحث، لكن هناك ظروفًا معينة، بل ومناهج بحثية (كمنهج دراسة الحالة)، وكذلك الدراسات الاستكشافية يمكن فيها إجراء الدراسة على عينات غير احتمالية؛ أي غير عشوائية، وبالتالي لا يمكن تعميم النتائج على المجتمع، ويكون التعامل مع هذه النتائج في حدود المفردات أو الحالات التي أجريت عليها الدراسة.

ثانيًا: حجم العينة

يقصد بحجم العينة عدد المفردات التي ستجرى عليها الدراسة، وليس هناك حجم ثابت يصلح لجميع الدراسات، فبعض البحوث تُجرى على بضعة أفراد، أو عشرات أو مئات أو ألوف الأفراد، وقد لوحظ في كثير من المؤلفات التي تناولت مسألة حجم العينة العديد من المغالطات فيما يتعلق بالحد الأدنى والحد الأقصى لحجم العينة؛ كالقول مثلًا بأن الحد الأدنى للعينة في الدراسات الوصفية يجب أن يكون (10%) من المجتمع الأصلي، فهذا غير صحيح بالمرة، ماذا نقول مثلًا إذا كان حجم المجتمع الأصلي عشرة ملايين نسمة مثلًا، فهل يكون حجم العينة مليون شخص؟ لا شك في أنه من الصعب، إن لم يكن من المستحيل إجراء دراسة على عينة بهذا الحجم الضخم، وبوجه عام، إن حجم العينة يتم تحديده على ضوء أكثر من اعتبار أهمها:

1- طبيعة المجتمع:

فالمجتمع الذي ستسحب منه العينة له مواصفات من حيث العدد والتجانس، وهذا يساهم إلى حد كبير في تحديد حجم العينة وخصائصها أيضًا. فإذا كان المجتمع - على

سبيل المثال - هو التلاميذ مكفوفو البصـر، والبـالغ عـددهم مائـة تلميـذ وتلميـذة، فـإن الدراسـة قـد تشملهم جميعًا وقد تشمل خمسين أو ثلاثين منهم، أما إذا كان حجم المجتمع كبيرًا- مليون أو أكثر على سبيل المثال- فإن حجم العينة قد يكون بضع مئات من المفردات، أو ألف مفردة. ومن العوامل المؤثرة في تحديد حجم العينة مدى تجانس أفراد المجتمع فـيما يتصـل بموضـوع البحـث والمتغـيرات محتملـة التأثير فيه، فكلما كان المجتمع أقل تجانسًا (أي أكـثر تباينًـا)، يصـبح مـن الضروري اختيـار عينـة كبـيرة نسبيًا؛ حتى نضمن أن تشتمل العينة على عدد كافٍ من المفردات بما يضمن تمثيل المجتمع، أما إذا كـان التباين صغيرًا (بمعنى أن يكون المجتمع أكثر تجانسًا)، فإنه من الممكن تصغير حجم العينة. (مـن أمثلـة المجتمعات الأكثر تجانسًا (الأقل تباينًا) تلاميـذ صـف دراسي معـين في نمـط معـين مـن المـدارس ولـتكن مدارس اللغات، ومن أمثلة المجتمعات الأقل تجانسًا (الأكثر تباينًا) مجتمع المدن بفئاته المختلفة).

2- أدوات جمع البيانات:

إن أدوات جمع البيانات تتمثل في الاستبانات والاختبارات والمقاييس وغيرها، وفي بعض الدراسات يتم تطبيق أكثر من أداة على المفردة الواحدة، وفي دراسات أخرى يتم تطبيق أداة واحدة على المفردة. وكلما تعددت الأدوات (والتي ستطبق على كل فرد)، فإن ذلك يسـتدعي تصـغير حجـم العينـة؛ بحيـث تكون في حدود الوقت المطلوب والإمكانات المتاحة ، وهنا يتدخل تقدير الباحـث مـع الاسـتفادة بـآراء أهل الاختصاص. أما إذا كانت الدراسة تستخدم أداة واحدة، أو أدوات محـدودة فـيمكن زيـادة حجـم العينة، وفي كل الأحوال فإن حجم العينة يجب أن يتفق ومتطلبات الحصول على بيانـات كافيـة ونتـائج موثوق فيها.

3- التكلفة:

تتطلـب البحـوث تكلفـة ماليـة بـدرجات متفاوتـة لـزوم جمـع البيانـات، والسـفر والانتقالات، وشراء مستلزمات أو أجهـزة ... إلـخ، وكلـما زادت التكلفـة، كـان ذلـك أدعـى

إلى تقليل حجم العينة، والعكس صحيح؛ أي أن حجم العينة يتأثر- ضمن عوامل أخرى- من بينها الموارد المتاحة، فكلما كانت التكلفة مرتفعة بحيث يتعذر إجراء الدراسة على عينة من ألف مفردة مثلًا، يصبح تخفيض حجم العينة أمرًا ضروريًا لكي يكون في حدود الموارد المتاحة.

4- الزمن أو الوقت:

هناك بحوث يتعين إنجازها خلال فترة زمنية معينة، فإذا كان حجم العينة كبيرًا، فإن الدراسة لن تنجز في الوقت المحدد، وقد تستغرق الدراسة زمنًا طويلًا لإنجازها، وحينئذٍ يمكن أن تكون الظاهرة المدروسة قد تغيرت بصفة كلية أو جزئية فتقل القيمة العلمية للدراسة، وقد يتم إلغاء الدراسة إذا لم يتم إنجازها خلال فترة زمنية، فعلى سبيل المثال فإن العديد من الجامعات تلغي البحث إذا لم ينجزه الباحث في غضون سنوات محددة، وبالتالي فإن حجم العينة يتحدد - ضمن عوامل أخرى - في ضوء الفترة الزمنية التي يتعين خلالها إنجاز البحث.

5- طبيعة الدراسة وأهدافها:

إن العوامل المتعلقة بطبيعة البحث والهدف منه تؤثر بالتأكيد في حجم العينة، فقد يكون البحث ذا صفة استطلاعية، كأن يستهدف استطلاع آراء مشاهدي التليفزيون حول مستوى البرامج والمسلسلات، وهنا يمكن أن يكون حجم العينة بضع مئات أو حتى بضع ألوف، أما عندما يكون البحث تجريبيًا بهدف معرفة أثر طريقة جديدة في تقديم برامج التليفزيون،فإن حجم المجموعة الواحدة يتحدد حسب رؤية الباحث وتقييمه، فقد يختار الباحث مجموعة من ثلاثين فردًا كمجموعة تجريبية، ومجموعة أخرى مماثلة لها في العدد كمجموعة ضابطة. ويلاحظ هنا أن طبيعة الدراسة التجريبية تفرض أن يكون حجم العينة صغيرًا بما يتناسب مع المكان ومستلزمات التجربة.

6- احتمالات عدم الاستجابة:

يتأثر حجم العينة بتوقعات عدم الاستجابة، فكلما زادت التوقعات بأن يرفض بعض المبحوثين المشاركة في البحث، أو الإجابة على أسئلة معينة، يصبح من الضروري أخذ ذلك في الاعتبار وزيادة حجم العينة، فقد يقرر الباحث أن يجري دراسته على عينة قوامها (1000) مفردة، لكنه يرى أن من المحتمل أن يرفض بعض الأفراد المشاركة في البحث، أو يرفضون الإجابة على أسئلة معينة، ومن هنا يجري دراسته على 1050 مفردة؛ أي بزيادة قدرها خمسون مفردة عن العدد المقرر، وإذا أردنا إجراء بحث على 2000 حالة مثلًا وكان تقديرنا أن 20% من الحالات سوف لا تستجيب، فإنه من المناسب أن نبدأ بعينة حجمها 2400 حالة (علما بأن ذلك لا يقلل من أخطاء التحيز).

7- متغيرات الدراسة:

المتغير (Variable) هو الخاصية أو السلوك الذي نقيسه، فالنوع (ذكور & إناث) متغير، وكذلك المستوى التعليمي ومحل الإقامة وغير ذلك من المتغيرات التي تشكل الخصائص الديموجرافية للمبحوثين، وهناك متغيرات تعبر عن مظاهر سلوكية معينة على مستوى الأفعال (Actions) أو التفكير والاتجاهات والقيم والمعتقدات والآراء .... إلخ. وكلما تعددت متغيرات الدراسة يتعين زيادة حجم العينة، وتصبح هذه الزيادة أكثر إلحاحًا كلما تعددت التقسيمات الفرعية داخل كل متغير، فإذا كان متغير النوع مثلًا يتضمن تقسيمين فرعيين فقط (ذكور & إناث)، فإن هناك متغيرات تتضمن تقسيمات فرعية متعددة، كمتغير التعليم (أمي، قراءة وكتابة، الابتدائية، الإعدادية، الثانوية، فوق المتوسط، جامعي، ماجستير، دكتوراه). وهذه المتغيرات تأتي مع بعضها في جداول مزدوجة ،كأن تقتضي الدراسة رصد العلاقة بين متغير النوع ومتغير المستوى التعليمي، ففي هذه الحالة يتضمن الجدول 18 خلية (Cells)، فإذا لم تكن العينة كبيرة، فإن بعض الخلايا تصبح

فارغة (Empty Cells)؛ أي لا يوجد بها أرقام (فقد لا يوجد ذكور أو إناث في مستوى تعليمي معين)، مع ملاحظة أن هذا المثال يتعلق بمتغيرين فقط، فالواقع أن الدراسات العلمية تُعنى بتوزيع المفردات حسب أكثر من متغيرين(كتوزيع العينة حسب النوع والسن، والمستوى التعليمي مثلًا) فحسب النوع تتضمن العينة مجموعتين (ذكور & إناث) وحسب السن، فإن العينة قد تتضمن أربع مجموعات، أما حسب المستوى التعليمي، فإن العينة قد تتضمن تسع مجموعات؛ أي أن الجدول الذي يتضمن توزيع العينة حسب هذه المتغيرات الثلاثة يتكون من 72 خلية، وهذا يتطلب أن تكون العينة كبيرة بدرجة تكفي لوجود عدد مفردات في كل خلية بالجدول، فإذا كانت العينة صغيرة، فإن بعض الخلايا تكون فارغة تمامًا من الأرقام أو أن الأرقام الموجودة في خلايا الجدول تكون صغيرة؛ الأمر الذي لا يسمح بإجراء التحليلات الإحصائية المناسبة.

8- درجة الدقة المطلوبة:

من المعروف أنه كلما صغر حجم العينة، زاد تأثر النتائج بعامل الصدفة، وكلما زاد تأثير عامل الصدفة، انخفضت الثقة في دقة النتائج، ووفقًا لذلك، فإن حجم العينة يتعين أن يكون كبيرًا بدرجة كافية للتقليل من تأثير عامل الصدفة، وبالتالي زيادة الثقة في دقة النتائج. وفي المجالات العلمية التطبيقية (كالبحوث الطبية والهندسية والفيزيائية والكيميائية مثلًا) يشترط درجة دقة عالية جدّا، بمعنى خفض احتمال الشك أو عدم الثقة إلى أدنى حد، كأن يكون 0.0001 مثلًا، أما في العلوم الإنسانية، فإنه من الشائع استخدام مستوى المعنوية 0.05 أو 0.01 إن استخدام المعنوية 0.05 يقابله مستوى ثقة 95 %، بمعنى أننا إذا سحبنا 100 عينة عشوائية من المجتمع الأصلي، فإن ذلك يعني احتمال أن 95 عينة تكون دقيقة أو صحيحة مقابل خمس عينات تكون خاطئة، أما استخدام المعنوية 0.01 فيقابله مستوى ثقة 99 % بمعنى أننا إذا سحبنا 100 عينة من المجتمع الأصلي، فإن ذلك يعني احتمال أن 99 عينة تكون دقيقة أو صحيحة مقابل عينة واحدة تكون خاطئة

ولتقليل احتمال ظهور العينات الخاطئة يتعين زيادة مستوى الثقة، وهذا يعني حجم عينة أكبر، فعندما نحدد حجم العينة بمستوى الثقة 0.01، فإن العينة تكون أكبر بكثير مقارنة باستخدام مستوى الثقة 0.05

يتصل بالثقة في النتائج حدود الخطأ المسموح به زيادة أو نقصًا، فإذا تقرر أن يكون هذا الخطأ في أضيق الحدود، يتعين زيادة حجم العينة والعكس صحيح، فالباحث قد يقرر أنه عندما تكشف نتائج تحليل بيانات العينة عن أن 13% يقرؤون الصحف اليومية بانتظام، فإنه سيتعامل مع هذه النسبة على أنها تقل 2% أو تزيد 2%؛ أي أنها تتراوح ما بين 11% إلى 15%؛ أي أن هامش الخطأ المسموح به 2%. لكن الباحث قد يقرر أنه عندما تكشف نتائج تحليل العينة عن أن 13% يقرؤون الصحف اليومية بانتظام، فإنه سيتعامل مع هذه النسبة على أنها تقل 4% أو تزيد 4%؛ أي أن هامش الخطأ المسموح به 4%، بمعنى أن النسبة تتراوح بين 9% إلى 17%، في الحالة الأولى - عندما كان هامش الخطأ المسموح به صغيراً أو محدودًا- يتعين زيادة حجم العينة، أما في الحالة الثانية - عندما يكون هامش الخطأ المسموح به كبيرًا- فإنه يمكن تقليل حجم العينة. وبوجه عام،كلما أراد الباحث مستوى دقة عالية، يتعين اختيار عينة أكبر، وذلك مقارنة بما إذا اكتفى الباحث بمستوى دقة أقل.

في ضوء هذه الاعتبارات يتم تحديد الحجم المناسب للعينة، ويمكن أن يقتدي الباحث بأحجام العينات التي استخدمت في دراسات علمية مشابهة لدراسته، كما يمكنه الاستعانة بآراء المتخصصين. وهناك نقطة جوهرية يجب أن يعيها الباحثون وهم بصدد تحديد حجم العينة، فالعينة ذات الحجم الأكبر لا تعني بالضرورة أن تكون نتائجها موثوقًا فيها بدرجة أكبر مقارنة بنتائج مستمدة من عينة صغيرة الحجم، ذلك أن الثقة في نتائج العينة تتوقف على طريقة سحب العينة؛ أي أن يتم اختيار العينة بطريقة عشوائية صحيحة؛ بحيث تكون ممثلة للمجتمع وتعكس ما فيه من تباين.

ثالثًا: بعض الطرق الإحصائية لتقدير حجم العينة

هناك الكثير من الطرق الإحصائية لتقدير حجم العينة سواء من المجتمعات الصغيرة المحدودة، أو من المجتمعات الكبيرة التي يصل تعدادها إلى الألوف أو الملايين، كما أن الطرق الإحصائية المستخدمة في تقدير حجم العينة، ترتكز على معادلات ذات صيغ متعددة (بمعنى أن المعادلة الواحدة لها أكثر من صيغة) رغم أن أساسها الجبري واحد.

طريقة الاحتمال المقدر:

وتعرف بالطريقة العامة، وتقوم على المعادلة الآتية:

$$\text{د ق} = \text{د ع}\sqrt{\frac{\text{ح}(١-\text{ح})}{\text{ن}}}$$

حيث إن:

ن : حجم العينة المطلوبة

دع: الدرجة المعيارية المقابلة لمستوى الثقة. إذا استخدمنا مستوى الثقة 95%، فإن الدرجة المعيارية المقابلة له هي 1.96 وعندما يكون مستوى الثقة 0.95، فإن ذلك يعني أن درجة الشك 0.05 و تفترض نسبة 0.05 أنه إذا سحبنا مائة عينة من المجتمع، فإن هناك احتمالًا بأن خمسًا وتسعين عينة تكون دقيقة مقابل خمس عينات تكون غير دقيقة؛ أي أن العينة غير الدقيقة تظهر مرة كل عشرين مرة. أما إذا استخدمنا مستوى الثقة 99 %، فإن الدرجة المعيارية المقابلة له هي 2.58، وعندما يكون مستوى الثقة 0.99، فإن ذلك يعني أن درجة الشك 0.01 وتفترض نسبة 0.01 أنه إذا سحبنا مائة عينة من المجتمع، فإن هناك احتمالًا بأن تسعًا وتسعين عينة تكون دقيقة مقابل عينة واحدة تكون غير دقيقة؛ أي أن العينة غير الدقيقة تظهر مرة كل مائة مرة.

د ق: درجة الدقة، ويحددها الباحث بناء على أسس موضوعية وقناعاته الشخصية، كأن يقرر مثلًا بأنه يرتضي أن تكون درجة الدقة ± 5% بمعنى أنه إذا كشفت الدراسة عن

أن نسبة مشاهدة التليفزيون هي 75%، فإن هذه النسبة سنتعامل معها على أنها تقل 5% أو تزيد بمقدار 5%؛ أي أن النسبة ستكون في حدود من 70% إلى 80%.

ح : احتمال قيمة الخطأ المعياري للنسبة، وهذه القيمة تساوي 0.5 وبالتالي فإن (1- ح) تساوي 0.5؛ أي أن القيمة: ح (1- ح) في المعادلة السابقة يتم التعويض عنها بالقيمة (0.5× 0.5)، وذلك باعتبار أن النسبة (0.5 × 0.5) تعطي أكبر النتائج دقة قياسًا على أي نسبتين أخريين (مثال ذلك: (0.1 × 0.9) أو (0.2 × 0.8) أو (0.3× 0.7)، أو (0.4×0.6) ... إلخ. فإذا حددنا درجة الدقة بأنها ± 5% وحددنا مستوى الثقة بأنه 0.95 وبالتعويض في المعادلة المذكورة، يمكننا تحديد حجم العينة المطلوبة كالآتي:

$$\%5 = 1.96 \sqrt{\frac{\text{ح}(١-\text{ح})}{\text{ن}}}$$

$$\%5 = 1.96 \sqrt{\frac{٠,٥ \times ٠,٥}{\text{ن}}}$$

بالتخلص من الجذر التربيعي، وبالتعويض فإن:

$$0.0025 = \frac{٠,٥ \times ٠,٥ \times ٣,٨٤١٦}{\text{ن}}$$

$$0.0025 \text{ ن} = 0.5 \times 0.5 \times 3.8416$$

$$0.0025 \text{ ن} = 0.9604$$

$$\text{ن} = 0.9604 \div 0.0025 = 384.16$$

أي أن حجم العينة المطلوبة هو 384 مفردة، وذلك بمستوى ثقة 0.95 ودرجة دقة 0.05

ونلفت الانتباه هنا إلى أن الباحث يمكنه استخدام درجات دقة أخرى بشرط وجود مبررات منطقية لذلك، كأن يقرر - لأسباب علمية- أنه سيحدد حجم العينة بدرجة دقة

2.5% أو 3% أو 4% أو غير ذلك من درجات الدقة، فقد يقرر الباحث مثلًا أن درجة الدقة هي 4% ومستوى الثقة 99% وهنا يكون حجم العينة كالآتي:

$$\%4 = 2.58 \sqrt{\frac{\text{٠,٥} \times \text{٠,٥}}{\text{ن}}}$$

بالتخلص من الجذر التربيعي وإيجاد قيمة (ن) فإن:

$$0.0016 = \frac{\text{٠,٥} \times \text{٠,٥} \times \text{٦,٦٥٦٤}}{\text{ن}}$$

$$0.0016\ \text{ن} = 0.5 \times 0.5 \times 6.6564$$

$$0.0016\ \text{ن} = 1.6641$$

$$\text{ن} = 1.6641 \div 0.0016 = 1040$$

أي أن حجم العينة المطلوبة هو 1040 مفردة وذلك بدرجة دقة 4% بمستوى ثقة 99%

معادلة الرابطة الأمريكية للتربية:

عندما يكون المجتمع متجانسًا، فإن هناك معادلة اعتمدتها الرابطة الأمريكية للتربية لتحديد حجم العينة المناسب (n) وذلك بدرجة ثقة 0.95 إذا عرفنا عدد المجتمع (N)، وهذه المعادلة هي:

$$\text{ن} = \frac{(\text{د ع})^{\text{٢}} \times (\text{٠,٢٥}) \times \text{م}}{(\propto)^{\text{٢}} (\text{م - ١}) + (\text{د ع})^{\text{٢}} (\text{٠,٢٥})}$$

حيث إن:

ن = حجم العينة المطلوبة

د ع : الدرجة المعيارية المقابلة لمستوى الثقة 95% وهذه الدرجة تساوي 1.96

م: عدد المجتمع

$\propto$: قيمة ألفا Alpha وهي تساوي 0.05 (لأن مستوى الثقة 0.95)، أما فيما يخص القيمة 0.25، فهي تعني احتمال قيمة الخطأ المعياري للنسبة مقسومًا على 2(قيمة هذا الاحتمال تساوي 0.5 كما سبقت الإشارة).

كمثال للتوضيح نفرض أننا نريد سحب عينة عشوائية من تلاميذ الصف الأول بالمرحلة الابتدائية في المدارس الحكومية، وتفيد الإحصاءات الرسمية أن عدد هؤلاء التلاميذ هو 38128 ، هنا نلاحظ أن مجتمع البحث متجانس من حيث المرحلة العمرية والدراسية، ولتحديد حجم العينة المطلوبة، فإن مستوى الثقة 0.95 يعني أن (د ع) تساوي 1.96 وبالتالي، فإن $(د\ ع)^2$ تساوي 3.8416 أما قيمة $(\propto)^2$ فتساوي 0.0025 (وهذه القيمة ناتجة عن 0.05× 0.05)، وبموجب ذلك فإن حجم العينة يساوي :

$$380.3 = \frac{٣٨١٢٨ \times (٠,٢٥) \times (٣,٨٤١٦)}{(٠,٢٥)(٣,٨٤١٦) + (٣٨١٢٧)(٠,٠٠٢٥)}$$

أي أن حجم العينة المطلوبة هو 380 مفردة، وقد يتم توزيع مفردات العينة توزيعًا متساويًا أو متناسبًا حسب خصائص المجتمع ذات الدلالة (كالنوع والمنطقة الجغرافية... إلخ)، وتمتاز المعادلة المذكورة بأنها تأخذ في الاعتبار عدد المجتمع، ومن الطبيعي أن يختلف حجم العينة باختلاف عدد المجتمع.

طريقة تقدير الانحراف المعياري:

وهي لا تعتمد على النسبة (Percentage) في التعبير عن درجة الدقة المطلوبة، وإنما تعتمد على قيمة الفرق المسموح به بين النتائج التي سنحصل عليها من العينة بشأن موضوع الدراسة وبين واقع هذا الموضوع في المجتمع، ويرمز لهذا الفرق بالرمز E كاختصار لكلمة (Error)، وتتطلب استخدام أو تقدير الانحراف المعياري للظاهرة في المجتمع (σ^{s})، ويشيع استخدام هذه الطريقة إذا كان هدف البحث تقدير متوسط الظاهرة في المجتمع، كما أنها في جوهرها صورة أخرى للطريقة العامة، وتقوم على المعادلة الآتية:

$$n = \left\{ \frac{Z(\sigma^{s})}{E} \right\}^{2}$$

حيث:

n = حجم العينة المطلوبة

σ^{s} = القيمة التقديرية للانحراف المعياري في المجتمع

E = قيمة الخطأ المسموح به (بمعنى الفرق المسموح به بين قيمة الظاهرة كما تكشفها العينة وبين قيمة هذه الظاهرة في المجتمع)

Z = الدرجة المعيارية لمستوى الثقة، وهذه الدرجة تكون 1.96 إذا استخدمنا مستوى الثقة 0.95 بينما تكون 2.58 إذا استخدمنا مستوى الثقة 0.99

فإذا كان الهدف من البحث تقدير متوسط الظاهرة في المجتمع، فإنه من الضروري تحديد الفرق المسموح به بين المتوسط الذي سنحصل عليه من العينة ومتوسط الظاهرة في المجتمع؛ أي أن يكون هذا الفرق في حدود معينة، ومن الضروري كذلك تقدير الانحراف المعياري للمجتمع. إن الفرق المسموح به بين متوسط العينة ومتوسط المجتمع يقرره الباحث أو المستفيد من الدراسة كأن يقرر أن هذا الفرق يكون قيمة معينة بالزيادة أو بالنقصان (فترة الثقة)، أما فيما يخص الانحراف المعياري للمجتمع، فإنه يمكن معرفته من خلال بحوث سابقة أجريت على عينات من هذا المجتمع وتتضمن قيمة الانحراف المعياري، أو من خلال إجراء دراسة استطلاعية محدودة، أو الاستناد على أي أساس منطقي آخر يمكن الأخذ به (Barrow, 2001, pp124-125).

كمثال توضيحي، نفرض أن المسئولين بجريدة الأهرام يريدون حجم العينة المناسب لتقدير متوسط الوقت اليومي الذي يقضيه محررو الجريدة في استخدام الحاسوب في التحرير الصحفي؛ وبحيث يكون الفرق بين المتوسط الذي سنحصل عليه من العينة

والمتوسط الموجود في المجتمع (جميع محرري الجريدة) في حدود ربع ساعة (15 دقيقة) زيادة أو نقصًا، وبفرض أن هناك دراسة علمية توضح أن متوسط الوقت اليومي الذي يقضيه الصحفيون في استخدام الحاسوب هو 3.5 ساعة بانحراف معياري ساعة واحدة (60 دقيقة)، هنا يمكن اعتبار هذه القيمة (60 دقيقة) هي الانحراف المعياري للمجتمع، وبالتالي يكون لدينا قيمة الخطأ المعياري (وهي 15 دقيقة بالزيادة والنقصان)، كما أن لدينا تقديرًا للانحراف المعياري للمجتمع وهو 60 دقيقة، فإذا حددنا مستوى الثقة بأنه 95%، فإن الدرجة المعيارية المقابلة له تساوي 1.96، وباستخدام المعادلة المذكورة فإننا نصل إلى حجم العينة المطلوبة كالآتي:

$$\text{حجم العينة} = \left[\frac{١,٩٦(٦٠)}{١٥}\right]^{٢} = (7.84)^{2} = 61.46$$

أي أن حجم العينة المطلوبة هو 61 مفردة، وذلك بمستوى ثقة 0.95 هذا الحجم هو الذي يحقق الثقة (Confidence Interval) (من + 15 إلى – 15) دقيقة.

وتغطي المعادلة المذكورة مدى واسعًا من التطبيقات في التوصل إلى حجم العينة، فإذا أردنا - على سبيل المثال- تحديد حجم العينة المطلوبة لتقدير متوسط قيمة ما تنفقه الأسرة في شراء الصحف اليومية في السنة، فإننا نحدد مستوى الثقة المطلوب، وليكن 0.95 كما نحدد الفرق المسموح به بين متوسط العينة ومتوسط المجتمع، وليكن عشرة جنيهات زيادة أو نقصًا، بمعنى أن الفرق بين المتوسط الموجود في المجتمع والمتوسط الذي سنحصل عليه من العينة يكون في حدود عشرة جنيهات زيادة أو نقصًا. أما العامل الأخير والأساسي الذي لا بد من تقديره فهو الانحراف المعياري لمتوسط الإنفاق السنوي على شراء الصحف اليومية في المجتمع (ويتم تقدير الانحراف المعياري في المجتمع بناء على دراسات سابقة أو بناء على عينة محدودة يتم اختيارها بعناية)، لنفرض أنه تم تقدير قيمة الانحراف المعياري بأنها (150) جنيهًا، بموجب ذلك يكون لدينا : درجة الثقة المطلوبة،

حدود الفرق المسموح به، تقدير للانحراف المعياري في المجتمع ، وهنا يمكننا تطبيق المعادلة السابقة كالآتي:

$$\text{حجم العينة} = \left[\frac{١,٩٦(١٥٠)}{١٠}\right]^{٢} = 864.36$$

أي أن حجم العينة المطلوبة يكون 864 مفردة وذلك عند مستوى الثقة 0.95 بفارق ± 150 ، أما إذا استخدمنا مستوى ثقة أعلى (99%) مع بقاء قيمة الفرق المسموح به على ما هي عليه، فإن حجم العينة سوف يزداد بدرجة ملحوظة، وقد يكون لزيادة حجم العينة مبررات موضوعية أو علمية، ففي هذه الحالة فإن حجم العينة يكون:

$$\text{حجم العينة} = \left[\frac{٢,٥٨(١٥٠)}{١٠}\right]^{٢} = 1497.69$$

لقد ارتفع حجم العينة إلى 1498 مفردة، وذلك عند استخدام مستوى الثقة 0.99 وهنا قد يجد الباحث أن حجم العينة كبير، ويمكنه تقليل حجم العينة بحيث يكون في حدود التكلفة والوقت المتاح. إن تقليل حجم العينة في هذه الحالة يكون إما من خلال تخفيض درجة الثقة، أو من خلال زيادة قيمة الفرق المسموح به، كأن يكون هذا الفرق 15 جنيهًا بدلًا من 10 جنيهات، بمعنى أن الفرق المسموح به بين المتوسط الذي سنحصل عليه من العينة والمتوسط الموجود في المجتمع يكون مداه خمسة عشر جنيهًا زيادة أو نقصًا، وهنا يتم تقدير حجم العينة كالآتي:

$$\text{حجم العينة} = \left[\frac{٢,٥٨(١٥٠)}{١٥}\right]^{٢} = 665.64$$

أي أن حجم العينة المطلوبة يكون 666 مفردة ، وذلك بفارق مسموح به ±15 ومستوى ثقة 99% (2.58 وحدات الانحراف المعياري)

تصحيح حجم العينة:

كثيرا ما يكون حجم المجتمع الذي ستسحب منه العينة معروفًا، كأن يكون عدد طلاب الجامعة، أو موظفي الحكومة أو موظفي شركات الاستثمار، أو سكان منطقة معينة....في هذه الحالة يتم استخدام معامل تصحيح إضافي لتقليل حجم العينة بعض الشيء، وذلك باستخدام المعادلة الآتية:

$$\text{ع م} = \frac{\text{ع}_{1}}{1 + \frac{(\text{ع}_{1} - 1)}{\text{ن}}}$$

حيث إن :

ع م: حجم العينة المعدلة (بعد استخدام معامل التصحيح الإضافي)

ع 1: حجم العينة الذي حصلنا عليه بموجب إحدى الطرق الإحصائية

(ولنفرض أنه 666 مفردة)

ن: عدد المجتمع

فالجزء الأسفل (المقام) في المعادلة المذكورة هو: حجم العينة الذي حصلنا عليه مطروحًا منه واحد صحيح مع القسمة على عدد المجتمع؛ أي: (666-1)÷ ن، ويضاف ناتج هذه القسمة إلى واحد صحيح؛ أي : 1+ (666-1) ÷ ن فيكون لدينا قيمة معينة، وبقسمة حجم العينة الأصلي على هذه القيمة نحصل على حجم العينة المعدل (ع م)، فإذا افترضنا أن حجم العينة الذي حصلنا عليه هو 666 وكان عدد المجتمع الأصلي هو (37000) مفردة، فإن حجم العينة المعدل يكون:

$$\text{ع م} = \frac{666}{1 + \frac{(666 - 1)}{37000}} = 653$$

أي أن حجم العينة المعدل أو المصحح يساوي 653 مفردة، وعلى الرغم من ذلك، فقد يتضح للباحث أن حجم العينة المعدل لا يقل كثيرًا عن حجم العينة الأصلي، خاصة إذا كان تعداد المجتمع كبيرًا (بالملايين مثلًا)، وكان حجم العينة الأصلي 400 مفردة أو أقل.

استخدام الجداول الإحصائية:

وضع الإحصائيون جداول لتحديد الحد الأدنى لحجم العينة المناسب إذا علمنا حجم المجتمع، ومن تلك الجداول هذا الجدول الذي يوضح الحد الأدنى لحجم العينة عند مستوى الثقة 0.95

المجتمع	العينة	المجتمع	العينة	المجتمع	العينة
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346
85	70	440	205	4000	351
90	73	460	210	4500	354
95	76	480	214	5000	357
100	80	500	217	6000	361
110	86	550	228	7000	364
120	92	600	234	8000	367
130	97	650	242	9000	368
140	103	700	248	10000	370
150	108	750	254	15000	375
160	113	800	260	20000	377
170	118	850	265	30000	379
180	123	900	269	40000	380
190	127	950	274	50000	381
200	132	1000	278	75000	382
210	136	1100	285	1000000 ∞	384

مأخوذ عن :

Krejcie, Robert V. and Daryle W. Morgan. "Determining Sample Size for Research Activities". Educational and Psychological Measurements, 30 (Autumn 1970), p. 608.

وننوه هنا إلى أن أحجام العينات حسب هذا الجدول محددة بمستوى الثقة 0.95 ويتضح من الجدول - على سبيل المثال- أنه إذا كان مجمل المجتمع عشر مفردات، فإن حجم العينة يكون عشر مفردات (أي أنه يتعين دراسة كل المجتمع)، ويوضح الجدول أيضًا أنه إذا كان مجمل المجتمع 210، فإن حجم العينة المناسب يكون 136، أما إذا كان مجمل المجتمع 1200، فإن العينة تكون 291 مفردة، وإذا كان مجمل المجتمع 75000، فإن العينة تكون 382 مفردة.

أما إذا كان مجمل المجتمع مليون أو أكثر، فإن حجم العينة يكون 384 مفردة... وهكذا، وكما هو واضح من الجدول، فإنه لا يتضمن جميع الأرقام التي تعبر عن مجمل المجتمع، فقد يكون مجمل المجتمع مثلًا 2458 وفي هذه الحالة يتم تحديد حجم العينة بأقرب قيمة لمجمل المجتمع موضحة بالجدول؛ حيث يتضح أن القيمة (2458) تقع بين 2400 و2600، لكنها أقرب إلى 2400 وعندما يكون مجمل المجتمع 2400، فإن الجدول يوضح أن حجم العينة يكون 331، وعلى ذلك فإن حجم العينة المطلوبة يكون 331 إذا كان المجتمع 2458، لكن إذا افترضنا أن لدينا مجتمعًا مجمله 2581، فإن أقرب قيمة مناظرة له بالجدول هو 2600، وبالتالي- حسب الجدول- يكون حجم العينة 335 (وهو حجم العينة إذا كان مجمل المجتمع 2600).

وإذا كان الجدول السابق يوضح الحد الأدنى لحجم العينة عند مستوى الثقة 0.95، فإن الباحث قد يرغب في زيادة مستوى الثقة، وفي هذه الحالة يمكن الاستعانة بالجدول الآتي والذي يوضح الحد الأدنى لحجم العينة عند مستويات معنوية مختلفة إذا علمنا عدد المجتمع:

المجتمع	درجة الثقة				
	0.95	**0.96**	**0.97**	**0.98**	**0.99**
200	132	150	168	185	196
300	166	200	234	276	289
400	196	240	291	343	384
500	217	273	340	414	475
600	234	300	384	480	565
700	248	323	423	542	652
800	260	343	457	600	738
900	269	360	488	655	823
1000	278	375	516	706	906
2000	322	462	696	1091	1655
3000	341	500	787	1334	2286
4000	350	522	842	1500	2824
5000	357	536	869	1622	3288
6000	361	546	906	1715	3693
7000	364	553	926	1788	4049
8000	367	558	942	1847	4364
9000	368	563	954	1895	4646
10000	370	566	964	1936	4899
15000	375	577	996	2070	5855
20000	377	583	1013	2144	4688
25000	378	586	1023	2191	6938
30000	379	588	1030	2223	7275
40000	381	591	1039	2265	7745
50000	381	593	1045	2291	8056
75000	382	595	1052	2327	8514
100000	383	597	1056	2354	8762
500000	384	600	1065	2390	9423
∝ 1000000	384	600	1066	2395	9513

ويعتبر هذا الجدول مهمّا جدّا للباحثين، فهو يوضح مثلًا أن المجتمع إذا كان تعداده مليونًا أو أكثر، فإن الحد الأدنى لحجم العينة يكون 384 مفردة عند مستوى الثقة 0.95، يزداد إلى 600 مفردة عند مستوى الثقة 0.96، ثم إلى 1066 عند مستوى الثقة 0.97 ، أما عند مستوى الثقة 0.98، فإن حجم العينة يكون 2395 يرتفع إلى 9513 مفردة، إذا أردنا أن يكون مستوى الثقة 0.99 وكما سبقت الإشارة، فإن الجدول لا يتضمن جميع الأرقام التي تعبر عن مجمل المجتمع، فالباحث الذي يجري دراسته على طلاب الصف الأول الثانوي في المدارس الحكومية قد يجد أن عدد هؤلاء الطلاب 67132 طالبًا وطالبة، فإذا أراد الباحث تحديد حجم العينة المناسب بمستوى ثقة 0.98، فإن الجدول يوضح حجم العينة يكون 2327 مفردة. لقد تم تحديد حجم العينة بأقرب قيمة لمجمل المجتمع موضحة بالجدول؛ حيث يتضح أن القيمة (67132) تقع بين 50000 و75000، لكنها أقرب إلى 75000 وعندما يكون مجمل المجتمع 75000، فإن الجدول يوضح أن حجم العينة يكون 2327 مفردة وذلك عند مستوى الثقة 0.98 وهكذا نتعامل مع عدد المجتمع بأقرب قيمة مناظرة له في الجدول عند مستوى الثقة المطلوب فنحصل على الحد الأدنى لحجم العينة.

المبحث الثاني

أنواع العينات

هناك الكثير من أنماط العينات، ويلاحظ أن النمط الواحد له أكثر من مسمى، غير أن ما يعنينا في هذه الجزئية توضيح أهم العينات العشوائية والعينات غير العشوائية بصرف النظر عن تعدد المسميات:

أولًا: العينات العشوائية

العينة العشوائية (Random Sample): هي اختيار المفردات دون تدخل من الباحث بقصد أو تعمد أن تتضمن العينة مفردات معينة وعدم تضمنها مفردات أخرى؛ ولذلك

تسمى بالعينة غير المتحيزة، أو العينة الاحتمالية(Probability Samples)، وعندما تتوافر في العينـة صـفة العشـوائية، فإنـه يمكـن تعمـيم نتائجهـا عـلى المجتمـع. وتجـدر الإشـارة إلى أن مصـطلح العشـوائية (Randomness) في اختيار العينة لا يعني الفوضى أو عدم النظام كما قد يفهـم مـن لفـظ (العشـوائية)، وإنما يعني أن جميع أفراد مجتمع البحث تتاح لهم فرص متساوية لأن يتم اختيارهم ضمن العينة.

فإتاحة الفرصة المتساوية لجميع المفردات لأن يتم اختيارها ضـمن العينـة هـو الأسـلوب الأمثـل في اختيار عينة البحث العلمي، وعـلى الـرغم مـن أن هـذا الأسـلوب ينطـوي عـلى أخطـاء تسـمى أخطـاء المعاينة، إلا أنها تكون في الحدود المقبولة، بمعنى أنها لا تؤثر في صدق تمثيل العينـة للمجتمـع، وبالتـالي تكون النتائج ذات قيمة علمية وقابلة للتعميم على المجتمع الذي سحبت منه العينة.

وتتطلب العينة العشـوائية (Random Sample) تحديـد مجتمـع الدراسـة، بمعنـى تحديـد المجتمـع الأصلي الذي ستسحب منه العينة، ومعرفة خصائص هذا المجتمع من حيث التوزيـع الجغـرافي، ومـدى التجانس في الخصائص الديموجرافية، تلك الخصائص التي تمثـل متغـيرات يمكـن أن تـؤثر في الظـاهرة أو الموضوع محل البحث، وقد يكون مجتمع الدراسـة عبـارة عـن: التلاميـذ في مرحلـة دراسـية معينـة، أو أولياء الأمور، أو الأطفال في سن ما قبل المدرسة، أو ربات البيوت، أو شباب الجامعـات، أو المـوظفين أو المطلقين، أو الأرامل، أو التلاميذ الموهوبين، أو ذوي الإعاقة... إلخ.

أما عن أنواع العينة الاحتمالية (العشوائية) فإن أهمها يتمثل في: العينة العشوائية البسيطة، العينـة العشوائية المنتظمة، العينة العشوائية الطبقية، العينة العشوائية متعددة المراحـل. وفـيما يـلي توضـيح لكيفية اختيار هذه العينات:

(أ) العينة العشوائية البسيطة:

وهي العـينة التي يتم اختيار مفرداتها من قائمة شاملة لكل المفردات؛ وبحيث تتاح فرص متسـاوية لكل مفردة أن يتم اختيارها ضمن العينة بصرف النظر عن الاختـلافات بين المفـردات؛ أي دون تصـنيف هذه المفردات إلى طبقات أو مجمـوعات. من أمثـلة القوائم الشاملة: قائمة أسماء الطلاب، قائمة أسـماء الموظفين في جهة حكـومية، قائمة أسماء أعضاء المجلس المحلي، قائمة أسـماء أصـحاب المعـاشات، قائمـة أسماء سيدات الأعمال، قائمة أسماء المؤسسات الخدمية، قائمة أسماء الجمعيـات الأهليـة، قائمـة أسـماء أفلام الموسم، قائمة أسماء القرى... إلخ، فكل قائمة مـن هـذه القـوائم تتضـمن أسـماء المفـردات بأرقـام متسلسلة، وإذا لم تكن كذلك، فإن الباحث يقوم بإعدادها (كأن يقوم بحصر أسماء الجمعيـات الأهليـة وتدوين هذه الأسماء بأرقام متسلسلة في قائمة شاملة، أو حصر أسماء القرى، أو الشوارع، أو المنازل...).

ويتم اختيار العينة العشوائية البسيطة من القائمة التي تشمل جميع المفردات،فإذا افترضنا مثلًا أن لدينا فصلًا مدرسيّا يضم أربعين تلميذًا، ونريد اختيار عينة عشوائية بسيطة تتكون من 8 تلاميذ، فـإن هذا الاختيار يتم كالآتي:

- إعطاء رقم مسلسل لكل تلميذ، وذلك في قائمة مستقلة تتضمن أرقام وأسماء التلاميذ.
- كتابة كل رقم على ورقة أو بطاقة مستقلة، فيكون لدينا أربعون بطاقـة تحمـل أرقامًـا مسلسـلة من (1) إلى (40).
- خلط البطاقات جيدًا، واختيار ثمان بطاقات، فتكـون الأرقـام المدونـة عـلى هـذه البطاقـات هـي الأسماء التي تتضمنها العينة.

فإذا كانت هذه البطاقات تحمل أرقام 3، 5، 21، 13، 20، 18، 26، 32، فـإن ذلـك يعنـي أن العينـة ستجرى على التلاميذ الذين يحملون هذه الأرقام في قائمة الأسماء.

وتسمى هـذه الطريقـة بالعينـة العشـوائية البسـيطة (Simple Random Sample)، وقـد يـتم السحب مع الإعادة بمعنى أننا نختـار إحـدى البطاقـات ثـم نـدون مـا تحملـه مـن رقـم في ورقـة مستقلة، ثم تعاد البطاقة مرة أخرى ضـمن بـاقي البطاقـات، وهكـذا كلـما سـحبنا بطاقـة نـدون

الرقم الموجود بها ثم نعيدها ثانية ونواصل الاختيار حتى نستكمل العينة، وقد يحدث أن نختار بطاقة سبق اختيارها، وهنا تعاد ثانية ضمن البطاقات التي يتم السحب منها، وتستمر عملية الاختيـار عـلى ألا تتضـمن بطاقات سبق اختيارها،أما السحب مع عدم الإعادة، فيعني عدم إعادة البطاقة التي تـم اختيارهـا إلى بـاقي البطاقات التي يتم السحب منها.

وكثيرًا ما يستعين الباحثون بجداول الأعداد العشوائية في اختيار مفردات العينة العشوائية البسيطة، وهذه الجداول تم إعدادها وفق منهج علمي وتتضمن أرقامًا تم تنظيمها بطريقة خاصة؛ بحيث يمكـن استخدامها في اختيار مفردات العينات دون تحيز؛ إذ إنها تتيح فرصًا متساوية لجميع المفردات بأن يـتم اختيارها ضمن العينة، فإذا أردنا أن نسحب عينة قوامها 200 طالـب مـن بـين طـلاب إحـدى الكليـات والبـالغ إجمالهـم 500 طالـب، فـإن المطلـوب أن يكـون لـدينا قائمـة مسلسـلة بأسـماء هـؤلاء الطـلاب الخمسمائة، ثم نستخدم جدول الأرقام العشوائية بأن نضع أصبعنا بشكل عشوائي على أي رقـم في هـذا الجدول، وننظر إلى الأعداد الثلاثـة الأولى (نظـرًا لأن حجـم العينـة المطلوبـة 200 وهـذا الـرقم - 200- يتكون من ثلاثة أعداد) فإذا كان هذا الرقم الذي وضعنا إصبعـنا عليه يقل عن 500، فإنه يتم رصده في ورقة مستقلة، أما إذا كان يزيد على خمسمائة، فيـتم تجـاوزه والانتقـال إلى الـرقم الـذي يليـه، وهكـذا تستمر العملية حتى نحصل على 200 رقم من جدول الأرقام العشوائية، وبرصـد تلـك الأرقـام في ورقـة مستقلة، فإننا نختار أسماء الطلاب الذين يحملون نفس الأرقام في قائمة أسماء الطلاب.

على سبيل المثال إذا كانـت الأرقـام التـي تـم رصـدها مـن جـدول الأرقـام العشـوائية هـي (10097)، (37542)، (12807) فإن ذلك يعني اختيار الطلاب أرقام (100) ، (375)، (128) في قائمة الأسماء، ونلاحـظ هنا أننا أخذنا العدد ذا الأرقام الثلاثة على اليسار، غير أنه يمكن أن نأخذ العدد ذا الأرقام الثلاثـة مـن عـلى اليمين، بشـرط الالتزام بطريقة واحدة في اختيار العينة، والجدول التالي نمـوذج لأحـد الجـداول العشـوائية التي يمكن الاستعانة بها في الاختيار العشوائي لمفردات العينات التي تقل عن عشـرة آلاف مفردة:

جدول الأعداد العشوائية

10097	85017	84532	13618	23157	86952	02438	76520
37542	16719	82789	96041	05545	44109	05403	64894
08422	65842	27672	82186	14871	22115	86529	19645
99019	76875	20684	39187	38976	94324	43204	09376
12807	93640	39160	41453	97312	41548	93137	80157
66065	99478	70086	71265	11742	18226	29004	34072
31060	65119	26486	47353	43361	99436	42753	45571
85269	70322	21592	48233	93806	32584	21828	02051
63573	58133	41278	11697	49540	61777	67954	05325
73796	44655	81255	31133	36768	60452	38537	03529
98520	02295	13487	98662	07092	44673	61303	14905
11805	85035	54881	35587	43310	48897	48493	39808
83452	01197	86935	28021	61570	23350	65710	06288
88685	97907	19078	40646	31352	48625	44369	86507
99594	63268	96905	28797	57048	46359	74294	87517
65481	52841	59684	67411	09243	56092	84369	17468
80124	53722	71399	10916	07959	21225	13018	17727
74350	11434	51908	62171	93732	26958	02400	77402
69916	62375	99292	21177	72721	66995	07289	66252
09893	28337	20923	87929	61020	62841	31374	14225
91499	38631	79430	62421	97959	67422	69992	68479
80336	49172	16332	44670	35089	17691	89240	26940
44104	89232	57327	34679	62235	79655	81336	85157
12550	02844	15026	32439	58537	48274	81330	11100
63606	40387	65406	37920	08709	60623	02237	16505
61196	80240	44177	51171	08723	39323	05798	26457
15474	44910	99321	72173	56239	04595	10836	95270
94557	33663	86347	00926	44915	34823	51770	67897
42481	86430	19102	37420	41976	76559	24358	97344
23523	31379	68588	81675	15694	43438	36879	73208
04493	98086	32533	17767	14523	52494	24826	75246
00549	33185	04805	05431	94598	97654	16232	64051
35963	80951	68953	99634	81949	15307	00406	26898
59808	79752	02529	40200	73742	08391	49140	45427
46058	18633	99970	67348	49329	95236	32537	01390
32179	74029	74717	17674	90446	00597	45240	87379
69234	54178	10805	35635	45266	61406	41941	20117
19565	11664	77602	99817	28573	41430	96582	01758
45155	48324	32135	26803	16213	14938	71961	19476
94864	69074	45753	20505	78317	31994	98145	36168

وبوجه عام، فإن العينة العشوائية البسيطة- وإن كانت تتيح فرصًا متساوية لجميع مفردات المجتمع لاختيارها ضمن العينة- إلا أن تنفيذها يصبح صعبًا إذا كان حجم المجتمع كبيرًا جدًا (بالألوف أو بالملايين مثلًا). من جهة أخرى، فإن العينة العشوائية البسيطة لا تضمن تمثيل الاختلافات في المجتمع، لنفرض أن لدينا فصلًا مدرسيًا يضم 20 من الذكور، مقابل 20 من الإناث، فإذا كنا نريد سحب عينة مكونة من 8 مفردات، فإنه يفترض أن تضم 4 ذكور مقابل 4 إناث، غير أن ذلك قد يتحقق وقد لا يتحقق إذا استخدمنا العينة العشوائية البسيطة، فقد تأتي العينة موزعة بالتساوي بين الجنسين، كما قد تأتي متضمنة 5 ذكور مقابل 3 إناث، أو 6 إناث مقابل 2 ذكور أو غير ذلك من التوزيعات التي تجعل تركيب العينة يختلف عن تركيب الأصل الذي سحبت منه. كما أن العينة العشوائية البسيطة تصبح غير ممكنة عندما يكون مجتمع البحث كبيرًا وموزعًا على مساحة جغرافية شاسعة، ولذلك يلجأ الباحثون إلى جداول الأرقام العشوائية، كما يلجأون إلى العينة الطبقية عندما يتطلب البحث تمثيل العينة للمجتمع حسب الطبقات أو الفئات بما يعكس المتغيرات المطلوب دراسة تأثيرها في الظاهرة محل الدراسة.

(ب) العينة العشوائية المنتظمة:

العينة العشوائية المنتظمة (Systematic Random Sample): هي تلك العينة التي تقوم على العشوائية والانتظام في اختيار المبحوثين، والعشوائية تكون في اختيار المفردة الأولى، بينما يكون الانتظام في تساوي المدى (Range) بين المفردات التي يتم اختيارها في العينة. لتوضيح ذلك، افترض أننا نريد سحب عينة عشوائية منتظمة قوامها ثمان مفردات من فصل مدرسي يضم أربعين تلميذًا. في هذه الحالة نعطي للتلاميذ أرقامًا متسلسلة في القائمة المتضمنة أسماءهم، فيكون لدينا أربعون رقمًا، من 1 إلى 40، ثم نحدد المدى (Range) بقسمة إجمالي عدد تلاميذ الفصل على حجم العينة المطلوبة؛ أي:

$$\frac{٤٠}{٨} = 5$$

أي أن قيمة المدى تساوي خمسة، بعد ذلك نضع أرقامًا مسلسلة من 1 إلى 5؛ بحيث يكتب كل رقم في ورقة مستقلة، وتخلط هذه الأوراق الخمس، ويتم اختيار إحدى الأوراق منها. افترض أننا اخترنا الورقة التي تحمل الرقم 3، فيكون التلميذ الأول في العينة هو التلميذ رقم 3 بالقائمة الأصلية (المتضمنة أسماء جميع التلاميذ بأرقام متسلسلة) ثم يضاف الرقم 5، فيكون التلميذ الثاني هو 8، وهكذا إلى أن نصل إلى العدد المطلوب اختياره في العينة، وبالتالي تضم العينة التلاميذ أرقام: 3، 8، 13، 18، 23، 28، 33، 38 (لاحظ أن بين كل مفردة أخرى عدد متساوٍ من المفردات، هذا العدد المتساوي هو المدى، ويبلغ في مثالنا هذا خمس مفردات).

وكما هو واضح، فإن مفردات العينة العشوائية المنتظمة تتحدد بمجرد اختيار المفردة الأولى، كما يتضح أن العينة العشوائية المنتظمة تتطلب وجود قائمة شاملة وبأرقام متسلسلة لجميع مفردات المجتمع (وهذا لا يتحقق في أحيان كثيرة)، كما أن كفاءة العينة العشوائية المنتظمة في تمثيل المجتمع تتطلب أن تكون قائمة مفردات المجتمع مرتبة عشوائيًا كأن تكون الأسماء مدونة في تلك القائمة بصرف النظر عن الجنس (ذكور & إناث، أو أن تكون مرتبة أبجديًا... إلخ)، فإذا كانت المفردات في القائمة الأصلية مرتبة ترتيبًا مقصودًا، فإن ذلك يجعل المفردات التي يتم اختيارها لا تعبر عن الأسماء الموجودة في القائمة؛ أي استبعاد بعضها فلا يتم اختيارها ضمن العينة رغم أنها موجودة في المجتمع.

(ج) العينة العشوائية الطبقية:

العينة العشوائية الطبقية (Stratified Random Sample): هي تلك العينة التي يتم اختيار مفرداتها عشوائيًا من الطبقات أو الفئات التي يتكون منها مجتمع البحث، ويتم اختيار تلك العينة من خلال تقسيم المجتمع إلى فئات، كل فئة تضم المفردات التي تشترك في صفة معينة. ومن بين كل فئة يتم السحب العشوائي للمفردات المطلوبة. ولتوضيح هذه الطريقة ببساطة، افترض أنه في مثالنا السابق كان عدد تلاميذ الفصل مائة تلميذ وتلميذة،

منهم (70) ذكور، مقابل (30) إناث. إن اختيار عينة طبقية من عشر مفردات مثلًا يعني أن تتم عملية الاختيار عشوائيّا على مستوى الذكور على حدة، ثم على مستوى الإناث؛ بحيث تتضمن العينة مفردات من الجنسين، والاختيار هنا يمكن أن يتم بالأسلوب العشوائي البسيط، أو بالأسلوب العشوائي المنتظم (سبقت الإشارة إلى ذلك)، والعينة التي يتم اختيارها يمكن توزيعها بالتساوي بين الجنسين، بمعنى خمسة ذكور مقابل خمس إناث، كما يمكن توزيعها بالتناسب بين الجنسين؛ أي أن يكون عدد المفردات التي يتم اختيارها من كل جنس يتناسب مع إجمالي عدده الأصلي، وبالتالي فإن العينة التي تضم عشر مفردات تعني أن يتم سحب سبع مفردات من الذكور، وثلاث مفردات من الإناث.

إن هذا المثال لمجرد التبسيط؛ لأن تقسيم التلاميذ إلى فئات، يمكن أن يتم على أساس مستوى التحصيل، الذكاء، السن، منطقة الإقامة ... إلخ، كما أن المجتمع قد يكون بالآلاف أو حتى بالملايين، افترض على سبيل المثال، أن إجمالي عدد السكان في مدينة معينة هو 139681، منهم 96241 من الذكور، مقابل 43440 من الإناث؛ أي أن السكان يتوزعون بين الذكور بنسبة 69%، والإناث بنسبة 31%، وأننا نريد سحب عينة عشوائية طبقية حجمها (500) مفردة من الجنسين لإجراء دراسة عن قراءة الصحف، في هذه الحالة، فإن عدد المفردات التي يتعين سحبها من فئة الذكور يكون كالآتي:

$$\frac{\text{عدد الذكور}}{\text{إجمالي السكان}} \times \text{حجم العينة}$$

$$= \frac{٩٦٢٤١}{١٣٩٦٨١} \times 500 = 344.5$$

أي أن إجمالي عدد الذكور في العينة يكون (345) مفردة؛ ونظرًا لأن مجمل العينة 500 مفردة، فإن عدد الإناث في العينة يكون:

500 – 345 = 155 مفردة

ويمكن الحصول على عدد الإناث في العينة كالآتي:

$$\frac{\text{عدد الإناث}}{\text{إجمالي السكان}} \times \text{حجم العينة}$$

$$= \frac{٤٣٤٤٠}{١٣٩٦٨١} \times 500 = 155 \text{ مفردة تقريبًا}$$

أي أن العينة تتوزع بين الذكور بنسبة 69%، والإناث بنسبة 31%، وهي نسبة كل جنس في مجتمع البحث، فكأن تركيب العينة عدديًا يعكس تركيب المجتمع.

وإذا كان هذا يسمى بالتوزيع المتناسب، فإن هناك نوعًا آخر من التوزيع الطبقي للعينة يسمى التوزيع الأمثل، وهو يمكن الحصول عليه فقط عندما يكون الانحراف المعياري لكل فئة أو طبقة معروفًا. افترض أننا نريد سحب عينة قوامها 400 مفردة من مجتمع يضم ثلاث مناطق (منطقة ريفية، منطقة حضرية، منطقة بدوية)؛ وذلك لدراسة استخدام القنوات الفضائية الأجنبية في تلك المناطق، وكان سكان كل منطقة كالآتي:

- المنطقة الريفية 1851321
- المنطقة الحضرية 1623410
- المنطقة البدوية 1411273

أي أن إجمالي سكان هذه المناطق الثلاث هو (4886004) وقد تم اختيار 100 مفردة، وأجريت عليهم دراسة استطلاعية من سؤال واحد بشأن عدد الساعات التي يقضيها الشخص أسبوعيًا في مشاهدة القنوات الفضائية الأجنبية، وقد كشف تحليل استجابات المفحوصين عن أن الانحرافات المعيارية لمتوسط ساعات المشاهدة الأسبوعية لكل مجموعة كالآتي:

- المنطقة الريفية: الانحراف المعياري = 2.6

- المنطقة الحضرية: الانحراف المعياري = 1.7
- المنطقة البدويـة: الانحراف المعياري = 3.4

فإذا كان المطلوب هو سحب عينة قوامها 400 مفردة، فإن العدد الأمثل الـذي يتعـين سـحبه مـن كل مجـموعة يمكن الحصـول عليه بموجب المعادلة:

$$\frac{\text{عدد المجموعة المعنية} \times \text{انحرافها المعياري}}{\text{عدد مفردات كل مجموعة} \times \text{انحرافها المعياري}} \times \text{حجم العينة}$$

أي أن المقام يظل ثابتًا (فالمقام عبارة عـن عـدد مفـردات كـل مجموعـة مـن المجموعـات الـثلاث مضروبًا في انحرافها المعياري)، أما البسط فهو عبارة عن عدد مفردات المجموعة المطلوب تحديـد عـدد مفرداتها مضـروبًا في الانحراف المعياري لتلك المجموعة، وباستخراج قيمة المقـام (الثابـت)، فـإن هـذه القيمة تساوي:

$$(1.7 \times 1623410) + (2.6 \times 1851321)$$

$$12371559.8 = (3.4 \times 1411273) +$$

أي أن قيمة المقام أو الثابت هي 12371559.8

بناء على ذلك، فإن عدد المفردات التي يتم سحبها من مجموعة الريف

$$156 = 400 \times \frac{٢,٦ \times ١٨٥١٣٢١}{١٢٣٧١٥٥٩,٨} =$$

بينما يكون عدد المفردات التي يتعين سحبها من مجموعة الحضر

$$89 = 400 \times \frac{١,٧ \times ١٦٢٣٤١٠}{١٢٣٧١٥٥٩,٨} =$$

أما عدد المفردات التي يتعين سحبها من مجموعة البدو فهو:

$$155 = 400 \times \frac{٣,٤ \times ١٤١١٢٧٣}{١٢٣٧١٥٥٩,٨}$$

أي أنه حسب التوزيع الأمثل، فإن العينة تتوزع بين مجموعة الريف بواقع 156 مفردة، ومجموعة الحضر بواقع 89 مفردة، ومجموعة البدو بواقع 155 مفردة، وهذا يختلف عما إذا كنا قد اتبعنا طريقة التوزيع المتناسب.

فحسب التوزيع المتناسب، نجد أن مجموعة الريف تضم 1851321؛ أي أنها تشكل 38% تقريبًا من مجمل المناطق. فإذا كان حجم العينة الكلية 400 مفردة، فإن التوزيع المتناسب يعني أن يكون عدد المفردات المسحوبة من مجموعة الريف هو 152 (بمعنى 400 × 38 ÷ 100 = 152) ، أما فيما يخص مجموعة الحضر، فإنها تشكل 33.2% من مجمل مناطق البحث، وبالتالي فإن التوزيع المتناسب يقضي بأن يكون عدد مفرداتها في العينة هو 133 (بمعنى 400 × 33.2 ÷ 100) وعلى مستوى مجموعة البدو، فإنها تشكل 28.9% من مجمل مجتمع البحث، وبالتالي، فإن التوزيع المتناسب يقضي بأن يكون عدد مفرداتها في العينة يساوي 115 (بمعنى 400 × 28.8 ÷ 100) هكذا يتضح أنه عند مقارنة توزيع مفردات العينة حسب أسلوب التوزيع الأمثل وأسلوب التوزيع المتناسب، فإن هناك اختلافًا يوضحه الجدول الآتي:

المجموعة	مجمل المجتمع	التوزيع المتناسب	التوزيع الأمثل
مجموعة الريف	1851321	152	156
مجموعة الحضر	1623410	133	89
مجموعة البدو	1411273	115	155
الإجمالي	**4886004**	**400**	**400**

من الواضح أنه حسب التوزيع المتناسب، فإن نسبة مفردات كل مجموعة في العينة تتساوى مع نسبة عدد هذه المجموعة في المجتمع، على سبيل المثال، فإن مجموعة الريف تضم 1851321؛ أي أنها تشكل قرابة 38% من الإجمالي البالغ 4886004، وفي الوقت

نفسه نجد أن عدد مفردات مجموعة الريف في العينة هو 152 مفردة؛ أي ما يعادل 38% مـن العينـة، المنطق نفسه مع اختلاف الأرقام فيما يخص مجموعتي الحضـر والبدو.

لكن عند استخدام أسلوب التوزيع الأمثل اختلف الأمر بعض الشيء، فقـد انخفـض عـدد مفـردات مجموعة الحضر في العينة ليصبح 89 مفردة فقط (بعد أن كان 133 مفردة حسب التوزيع المتناسـب)، هذا الانخفاض يفسر بأن مجموعة الحضر هي أكثر المجموعات تجانسًا؛ حيث تنخفض قيمة الانحـراف المعياري إلى 1.7 (ومن المعروف أنه يمكن تقليل حجم العينـة كلـما كـان المجتمـع أكـثر تجانسًـا)، أمـا مجموعة البدو، فإنها أكثر مجموعات العينة تباينًا؛ حيث ترتفع قيمة الانحراف المعيـاري إلى 3.4. ومـن هنا ارتفع عدد مفرداتها من 115 مفردة حسـب أسـلوب التوزيـع المتناسـب إلى 155 مفـردة حسـب التوزيع الأمثل (ومن المعروف أنه يتعين زيادة حجم العينة كلما المجتمع أكثر تباينًا).

أما مجموعة الريف، فقد كان عدد مفرداتها في العينـة 152، حسـب التوزيـع المتناسـب، ارتفـع إلى 156 حسب التوزيع الأمثل، وذلك يعزى أيضًا إلى التباين، فإذا كانت مجموعـة الريـف أقـل تباينًـا عـن مجموعة البدو، إلا أنها أكثر تباينًا من مجموعة الحضر.

بوجه عام، فإن التوزيع الأمثل هو الأشد دقة عند اختيار العينات؛ لأنه يأخذ تباين المجتمع في الاعتبار، لكن المشكلة في تطبيق هذا الأسلوب تتمثـل في ضرورة المعرفـة المسـبقة بقيمـة الانحـراف المعيـاري لكـل مجموعة أو فئة أو طبقة، وهذه المعرفة نادرًا ما تتحقق خاصة عندما يكون عدد المجموعات كبيرًا، وعـلى الرغم من إمكانية تقدير الانحراف المعياري للمجموعات بشأن المتغـيرات التـي يـتم عـلى أساسـها اختيـار العينة، إلا أن هذا التقدير عرضة للخطأ في كثير من الأحيان، وعادة يستخدم البـاحثون التوزيـع المتناسـب، أو المتساوي إلا في الحالات التي يكون الانحراف المعياري مقدرًا وفق معطيات قوية.

(د) العينة العشوائية متعددة المراحل:

وتسمى أحيانًا بالعينة العنقودية (Cluster Random Sample): وهي العينـة التـي يـتم اختيار مفرداتها عشوائيًا على عدة مراحل، ويلجأ البـاحـثون إلى هذه العينة عنـدما يكـون

حجم المجتمع كبيرًا، وتنتشر المفردات على مساحة جغرافية واسعة، وليس هناك قوائم شاملة للمفردات. افترض أنك بصدد إجراء دراسة عن اتجاهات الجمهور في مدينة معينة نحو برامج التلفاز، وأن تلك المدينة تضم عشر مناطق، في هذه الحالة يمكن اختيار ثلاث مناطق (أو أكثر أو أقل) بإحدى الطرق العشوائية، ومن بين كل منطقة تم اختيارها يتم اختيار شارع بشكل عشوائي، ويتم رصد أرقام البيوت الموجودة في كل شارع، ومن هذه البيوت يتم اختيار العشوائي للعدد المطلوب، ويتم إجراء المقابلات مع ساكني البيوت التي تم اختيارها، كأن يتم البحث مع مفردة من كل بيت أو وحدة سكنية سواء كانت منزلًا مستقلّا أو شقة، وقد تتم هذه العملية على مستوى جميع المدن في الدولة؛ بحيث تتعدد مراحل الاختيار لتصل في النهاية إلى شارع واحد من كل مدينة.

من الواضح أن ذلك يتطلب معرفة الباحث جيدًا بالمناطق والشوارع الموجودة بكل مدينة، وأن يتم الاختيار (عشوائيّا) على كل مستوى من هذه المستويات، فإذا افترضنا أن المدينة تتضمن عشر مناطق، وأننا نريد اختيار منطقة واحدة، في هذه الحالة يدون اسم كل منطقة في ورقة مستقلة، وتخلط الأوراق جيدًا، وتسحب واحدة منها فتكون المنطقة المدونة في تلك الورقة هي المنطقة التي ستجرى فيها الدراسة، المنطق نفسه عند اختيار الشوارع الموجودة في المدينة، ويتوزع عدد المفردات المسحوبة من كل منطقة توزيعًا متناسبًا مع العدد الكلي للمنازل الموجودة بها. أي أنه في العينة العنقودية يتم اختيار مناطق أو مجموعات (وليس أفرادًا)، وتسمى كل منطقة أو مجموعة عنقودًا، وبعد ذلك نختار الأفراد من كل عنقود. ويشترط أن يكون لكل تجمع أو عنقود نفس الخصائص، ومن أمثلة العناقيد التي يمكن استخدامها في العينات: الفصول الدراسية، المدارس، السجون، المستشفيات... إلخ.

وتمتاز العينة العنقودية بأنها قليلة التكلفة، وسهولة الحصول عليها، ويمكن إنجازها في وقت قصير نسبيّا، وعلى الرغم من ذلك، لا بد من وجود مبررات لاستخدامها. ولضمان كفاءة العينة العنقودية، يمكن الاختيار العشوائي - البسيط أو المنتظم - للمفردات، فإذا

كنا بصدد إجراء دراسة عن استخدام تلاميذ الصف الأول الابتدائي لأفلام الكارتون التليفزيونية، فإننا نختار عينة من فصول الصف الأول الابتدائي، ويتم الاختيار العشوائي لعينة من التلاميذ في الفصول التي تم اختيارها (أي أن الدراسة لا تتم على جميع التلاميذ في الفصول المختارة، وإنما على عينة عشوائية بسيطة أو منتظمة أو طبقية من هؤلاء التلاميذ).

ثانيًا: العينات غير العشوائية

إنها العينات غير الاحتمالية (Non-Probability Samples): وهي العينات التي لا تتاح في اختيارها فرص متساوية لجميع مفردات المجتمع أن يتم اختيارها ضمن العينة، وكمبدأ عام، فإن العينات غير الاحتمالية لا تعمم نتائجها المجتمع، ولا يمكن تقدير معالم المجتمع من إحصاءات تلك العينات، ومن أمثلة العينات غير العشوائية: العينة العارضة، والعينة المقصودة، وعينة الحصص:

(أ) العينة العارضة:

العينة العارضة (Accidental Sample): هي العينة المتاحة، بمعنى أن يجري الباحث دراسته على الأشخاص الذين يصادفهم، أو الذين تتاح مقابلتهم، كأن يذهب إلى التجمعات أو المؤسسات أو الأماكن التي يوجد فيها الأشخاص الذين يمكن أن يحصل منهم على المعلومات المطلوبة، ويجري المقابلة مع أي شخص يقابله، وقد تفيد نتائج العينة العارضة في حدود معينة، وإن كان من الصعب تعميم نتائجها على المجتمع. افترض أن الباحث يجري دراسة ميدانية على الجمهور لمعرفة علاقته ببرامج التليفزيون، وأن حجم العينة هو 1000 مفردة، فإذا أجرى الباحث مقابلات مع كل مَن يمكنه مقابلته إلى أن يستكمل العدد المطلوب دون مراعاة عشوائية الاختيار، فإنه في هذه الحالة يكون قد اعتمد على العينة العارضة أو المتاحة، ولا يمكن تعميم نتائج تلك العينة على المجتمع.

(ب) العينة المقصودة:

العينة المقصودة (Purposive Sample): تعني أن يتعمد الباحث أو يقصد إجراء الدراسة على فئة معينة، وقد يكون هذا التعمد لاعتبارات علمية، كوجود أدلة أو براهين مقبولة أو منطقية تؤكد أن هذه العينة تمثل المجتمع، في هذه الحالة تكون نتائج الدراسة مقبولة، وقد تكون العينة المقصودة مبررة لاعتبارات واقعية أو منطقية، كأن يتم إجراء دراسة على عينة من الذين حضروا أحد المؤتمرات العلمية لمعرفة رأيهم في تنظيم المؤتمر، أو أن يتم إجراء دراسة على الذين أصيبوا في حادث معين لمعرفة الآثار النفسية لهذا الحادث على هؤلاء المصابين.. فالعينة هنا مقصودة كما أنها منطقية. لكن هناك عينة مقصودة لاعتبارات غير علمية، كأن يتعمد الباحث إجراء الدراسة على عينة معينة بسبب سهولة الوصول إلى الأفراد، أو أنهم موجودون بالقرب من الباحث أو في المكان الذي يعمل فيه، في هذه الحالة تكون العينة مقصودة، لكن النتائج المستمدة منها لا تكون مقبولة علميّا، ولا يمكن تعميمها على المجتمع.

(ج) عينة الحصص:

يقصد بعينة الحصص (Quota Sample) أن تتضمن العينة عددًا من المفردات تنتمي إلى الفئات التي تشكل مجتمع البحث دون أي اعتبار آخر، فإذا كان المجتمع يضم الذكور والإناث، فإن العينة تتضمن مفردات من الذكور ومفردات أخرى من الإناث دون أن يتم اختيار المفردات بالطريقة العشوائية. افترض أنك تقوم بدراسة على مجتمع جامعي يضم مليون مفردة، يتوزعون بين طلاب التخصصات العلمية بواقع ستمائة ألف، والتخصصات الأدبية بواقع أربعمائة ألف، كما يتوزع هؤلاء الطلاب حسب الجنس بواقع خمسمائة ألف للذكور، مقابل خمسمائة ألف أيضًا للإناث، ويتساوى عدد الجنسين تقريبًا في التخصصات العلمية، وكذلك في التخصصات الأدبية، فإذا كان حجم العينة قد تحدد بألف مفردة، فإن عينة الحصص (Quota Sample) تتوزع كالآتي:

– ستمائة مفردة من التخصصات العلمية بواقع ثلاثمائة من الذكور، وثلاثمائة من الإناث.

– أربعمائة مفردة من التخصصات الأدبية بواقع مائتين من الذكور، ومائتين من الإناث.

ويتم إجراء المقابلات مع العدد المطلوب من كل فئة دون الالتزام بعشوائية الاختيار، فكل المطلوب هو أن تضم العينة «حصة» من كل فئة حسب وجودها في المجتمع. لا شك أن هذه الطريقة تتميز بالسهولة، كما يمكن أن يتوافر فيها بعض الدقة إذا كانت العينة تضم مجموعات من المفردات تعكس وزن تلك المجموعات في المجتمع على أساس أكبر عدد ممكن من المتغيرات التي يحتمل أن تؤثر في الظاهرة أو الموضوع محل البحث، وعلى الرغم من ذلك لا يمكن تعميم نتائج عينة الحصص على المجتمع لأنها تفتقد شرط عشوائية الاختيار، بمعنى أنه لم تتح الفرص المتساوية لكل مفردات المجتمع لتكون ضمن العينة.

أخيرًا، تجدر الإشارة إلى أن العينات - سواء كانت عشوائية أو غير عشوائية- يمكن تطبيقها على المفردات بوجه عام (وليس فقط على الأفراد)، فالمفردات (Subjects) قد تكون أشخاصًا، كما قد تكون مناطق أو شوارع، أو قري أو غير ذلك من الأماكن، كما أن المفردات قد تكون مؤسسات أو مدارس أو فصولًا مدرسية... إلخ. وفي مجال الاتصال، فإن المفردات قد تكون أخبارًا أو نشرات إخبارية، أو برامج أو مسلسلات أو قنوات إذاعية وتليفزيونية أو صحفًا، أو مقالات صحفية أو إعلانات... إلخ، وبوجه عام يمكن تطبيق إجراءات المعاينة أيّا كانت طبيعة المفردات.

المبحث الثالث

خطأ المعاينة

إن علاقة العينة بالمجتمع الذي سحبت منه تعتمد على حجم العينة وطريقة اختيارها. العينة الصحيحة هي التي تمثل المجتمع الذي سحبت منه تمثيلًا صادقًا، وتقترب العينة من الأصل الذي سحبت منه، كلما اقتربت مقاييسها الإحصائية من مقاييس ذلك الأصل، على سبيل المثال، نفرض أن لدينا مجتمعًا يضم 30 مليون شخص تتراوح أعمارهم ما بين 20 سنة إلى 90 سنة، وكان متوسط العمر لهذا المجتمع هو 65 سنة، ومن هذا المجتمع سحبنا عينة عشوائية قوامها 400 شخص، وحسبنا متوسط السن لهذه العينة، وباستخدام الطرق الإحصائية يمكن المقارنة بين متوسط الأعمار في العينة ومتوسط الأعمار في المجتمع، فإذا تبين أنه لا توجد فروق جوهرية بين المتوسطين، فإن هذا يعني أن متوسط أعمار العينة لا يختلف اختلافًا جوهريًا عن متوسط الأعمار في المجتمع الذي سحبت منه، وفي هذه الحالة تكون العينة ممثلة للمجتمع تمثيلًا صادقًا من حيث السن (ومن الضروري هنا أخذ متغيرات أخرى في الاعتبار مثل الجنس ومنطقة الإقامة... إلخ).

وعلى الرغم من ذلك، فإن العينة مهما كانت دقة اختيارها عشوائيًا لا تكون ممثلة للمجتمع مائة بالمائة، ولذلك هناك ما يعرف بخطأ المعاينة (Sampling Error) والذي ينعكس على المقاييس الإحصائية التي تجري على البيانات المأخوذة من العينة، بحيث تأتي مختلفة بعض الشيء عن المجتمع. إن ذلك ينطبق على كافة المقاييس الإحصائية، وسوف نوضح هنا ثلاثة أمثلة فقط هي: الخطأ المعياري للمتوسط، والخطأ المعياري للنسبة، والخطأ المعياري لمعامل الارتباط:

أولًا: الخطأ المعياري للمتوسط

يتم حساب الخطأ المعياري للمتوسط الحسابي للعينة $\overline{X}$ بمعلومية الانحراف المعياري وحجم العينة وذلك بموجب المعادلة:

$$\frac{ع}{\sqrt{ن}}$$

فإذا كان لدينا عينة عشوائية قوامها 400 شخص، وطبقنا عليها مقياس استخدام وسائل الإعلام، وتبين أن متوسط درجة العينة على هذا المقياس 8.7 درجة بانحراف معياري هو 3.6 فإن الخطأ المعياري للمتوسط (ع م) يكون:

$$ع\ م = \frac{٣,٦}{\sqrt{٤٠٠}} = 0.18$$

فإذا كانت قيمة الخطأ المعياري لمتوسط درجة العينة هي 0.18، فإن حدود هذا المتوسط هي:

الحد الأدنى = المتوسط - الخطأ المعياري = 8.7 - 0.18 = 8.52

الحد الأعلى = المتوسط + الخطأ المعياري = 8.7 + 0.18 = 8.88

أي أن القيمة العددية لمتوسط درجة العينة تمتد من 8.52 إلى 8.88

تقدير متوسط الظاهرة في المجتمع:

يمكن تقدير قيمة μ بمعنى متوسط الظاهرة في المجتمع وذلك بالاستفادة من المعطيات السابقة، فمع افتراض عشوائية العينة واعتدالية التوزيع التكراري، فإن المساحة المحصورة بين - ع و +ع تساوي 68% أي أن المساحة المتبقية تساوي 32% أي أن نسبة المساحة المحصورة إلى نسبة المساحة المتبقية هي 2 إلى 1 وهنا يمكن تقرير أن متوسط درجة المجتمع يقع في المدى (من 8.25 إلى 8.88) بدرجة ثقة 2 ودرجة شك 1، غير أننا يمكننا زيادة درجة الثقة إلى 95% وإنقاص درجة الشك إلى 5% إذا ضربنا الخطأ المعياري في 1.96 (لأن المساحة المعيارية المحصورة بين -1.96 إلى + 1.96 انحراف معياري تساوي 0.95 من المساحة الكلية للمنحنى الاعتدالي المعياري).

وفي المثال السابق، فإن متوسط درجة العينة هو 8.7 وبما أن الخطأ المعياري يساوي 0.18 فإن المدى الذي يقع فيه متوسط درجة المجتمع الذي سحبت منه العينة العشوائية المذكورة يساوي 8.7 $\pm$ (1.96× 0.18) أي

$$8.7 \pm 0.3528$$

أي أن المدى الذي يقع فيه متوسط المجتمع الذي سحبت من العينة المذكورة يمتد من 8.3472 إلى 9.0528 وذلك بدرجة ثقة 0.95 أما إذا استخدمنا درجة الثقة 0.99 فإن هذا المدى يكون:

$8.7 \pm (0.18 \times 2.58)$ أي:

$$8.7 \pm 0.4644$$

أي أن المدى الذي يقع فيه متوسط المجتمع الذي سحبت من العينة المذكورة يمتد من 8.2356 إلى 9.1644 بدرجة ثقة 0.99

ثانيًا: الخطأ المعياري للنسبة

يمكن حساب الخطأ المعياري للنسبة التي حصلنا عليها من عينة عشوائية؛ ومن ثمّ تقدير المدى الصحيح الذي تقع فيه تلك النسبة، فإذا كانت لدينا عينة عشوائية قوامها 500 مفردة، وتبين أن 37% من العينة يقرؤون الصحف اليومية، فإن لهذه النسبة خطأ معينًا، ويمكن حساب هذا الخطأ (الخطأ المعياري للنسبة) بموجب المعادلة:

$$\sqrt{\frac{\text{النسبة (1- النسبة)}}{\text{عدد الأفراد}}}$$

وبتطبيق هذه المعادلة على المثال المذكور يكون:

$$\sqrt{\frac{٠,٣٧\,(١ - ٠,٣٧)}{٥٠٠}}$$

أي

$$0.02 = \sqrt{\frac{٠,٣٧ \times ٠,٦٣}{٥٠٠}}$$

أي أنه إذا كانت نسبة قراءة الصحف اليومية من واقع بحث العينة هي 37% فإن الخطأ المعيـاري لهذه النسبة هو 2% وعلى ذلك فإن النسبة المصححة تتراوح مابين 35% إلى 39%

تقدير نسبة الظاهرة في المجتمع:

في المثال السابق، إذا كان الخطأ المعياري للنسبة هو 2% ومـع افـتراض عشـوائية العينـة واعتداليـة التوزيع التكراري، فإنه من واقـع النسـبة في العينـة (P) مكننـا تقـدير π (بـاي) معنـى نسـبة قـراءة الصحف اليومية في المجتمع سواء مستوى ثقة 0.95 أو 0.99 فإذا اتخذنا مستوى الثقة 0.95 فإن تقدير النسبة في المجتمع يكون:

$0.02 \times 1.96 = 0.04$ تقريبًا

وحيث إن النسبة المستمدة من العينة هي 37% فإن الحد الأدنى للنسبة المجتمع هو 0.37 - 0.04 = 33% بينما يكون الحد الأعلى للنسبة في المجتمع هو 0.37 + 0.04 = 41%

أي أن نسبة قراءة الصحف في المجتمع تقدر ما بين 33% إلى 41% مستوى ثقة 0.95

أما إذا اتخذنا مستوى الثقة 0.99 فإن تقدير النسبة في المجتمع يكون:

$0.02 \times 2.58 = 0.05$ تقريبًا

فإذا كانت نسبة قراءة الصحف اليومية من واقع بحث العينة هي 37% فإن الحد الأدنى للنسبة في المجتمع هو 0.37 – 0.05 = 32% ، أما الحد الأعلى للنسبة في المجتمع فهو 0.37 + 0.05 = 42%؛ أي أن نسبة قراءة الصحف في المجتمع تتراوح ما بين 32% إلى 42% وذلك بمستوى ثقة 0.99

ثالثًا: الخطأ المعياري لمعامل الارتباط

تتعدد صيغ حساب الخطأ المعياري لمعامل الارتباط، غير أن ما يعنينا في هذه الجزئية هو حساب الخطأ المعياري لمعامل الارتباط الذين يقع بين الحدين المتطرفين للصفر والواحد الصحيح، في هذه الحالة يتم حساب الخطأ المعياري لمعامل الارتباط بموجب المعادلة الآتية:

$$= \frac{١ - ر^{٢}}{\sqrt{ن}}$$

فإذا كان معامل الارتباط بين قراءة الصحف اليومية واستخدام الإنترنت = 0.6 وذلك من واقع دراسة على عينة عشوائية قوامها (400) مفردة، فإن الخطأ المعياري لمعامل الارتباط يكون:

$$ر = \frac{١ - (٠,٦)^{٢}}{\sqrt{ن}} = 0.032$$

أي أن معامل الارتباط يتراوح بين 0.6 ± 0.032

ويكون الحد الأدنى لمعامل الارتباط = 0.6 – 0.032 = 0.568

في حين أن الحد الأعلى لمعامل الارتباط = 0.6 + 0.032 = 0.632

أي أن قيمة معامل الارتباط بين قراءة الصحف اليومية واستخدام الإنترنت (في العينة) تتراوح ما بين (0.568) إلى (0.632)

تقدير معامل الارتباط بين الظاهرتين في المجتمع:

بموجب معرفة معامل الارتباط بين الظاهرتين من واقع عينة الدراسة، يمكن تقدير معامل الارتباط بين الظاهرتين في المجتمع الأصل، ففي المثال السابق تبين من دراسة العينة أن معامل الارتباط بين قراءة الصحف اليومية واستخدام الإنترنت هو (0.6)، وكان الخطأ المعياري لمعامل الارتباط هو (0.032) فإذا أراد الباحث تقدير هذا الارتباط في المجتمع الأصل، فإنه يتبع الإجراءات نفسها في تقدير النسبة أو المتوسط سواء بدرجة ثقة (0.95) أو بدرجة ثقة (0.99)

فإذا تم استخدام درجة الثقة (0.95)، يتم ضرب الخطأ المعياري في 1.96 أي:

$0.063 = 1.96 \times 0.032$

وبالتالي تتراوح قيمة معامل الارتباط ما بين 0.6± 0.063

أي أن الحد الأدنى لمعامل الارتباط = 0.6 – 0.063 = 0.537

في حين أن الحد الأعلى لمعامل الارتباط = 0.6 + 0.063 = 0.663

أي أن معامل الارتباط بين قراءة الصحف اليومية واستخدام الإنترنت في المجتمع الأصل تتراوح قيمته ما بين 0.537 إلى 0.663 بدرجة ثقة 0.95

أما إذا استخدمنا درجة الثقة 0.99، فإن

$0.083 = 2.58 \times 0.032$

وبالتالي تتراوح قيمة معامل الارتباط بين 0.6 ± 0.083

أي أن الحد الأدنى لمعامل الارتباط = 0.6 – 0.083 = 0.517

أما الحد الأعلى لمعامل الارتباط = 0.6 + 0.083 = 0.683

أي أن معامل الارتباط بين قراءة الصحف اليومية واستخدام الإنترنت في المجتمع الأصل تتراوح قيمته ما بين 0.517 إلى 0.683 بدرجة ثقة 0.99

خلاصة الفصل السادس

- تناول هذا الفصل شرحًا للعينات في البحث العلمي، مع التركيز على المنظور الإحصائي، وتتلخص الأفكار الأساسية فيما يلي:

- العينة (Sample) هي جزء من المجتمع تتوافر فيه خصائص هذا المجتمع، والحكمة من إجراء الدراسة على العينة تتمثل في أنه في كثير من الأحيان يستحيل إجراء الدراسة على المجتمع، وبالتالي يتم اختيار عينة محدودة ودراستها بهدف التوصل إلى نتائج يمكن تعميمها على المجتمع. ويصبح ذلك ممكنًا، إذا كانت خصائص العينة تمثل خصائص المجتمع من حيث أكبر عدد ممكن من المتغيرات، خاصة المتغيرات التي يحتمل أن تؤثر في الظاهرة محل البحث. والعينات في البحوث العلمية يفترض أن تكون ممثلة للمجتمع؛ حتى يمكن تعميم نتائج البحث.

- يقصد بحجم العينة عدد المفردات التي ستجرى عليها الدراسة، وليس هناك حجم ثابت يصلح لجميع الدراسات، ويتم تحديد حجم العينة على ضوء أكثر من اعتبار أهمها: طبيعة المجتمع (متجانس أو غير متجانس)، أدوات جمع البيانات، التكلفة، الزمن أو الوقت، طبيعة الدراسة وأهدافها، احتمالات عدم الاستجابة، متغيرات الدراسة، درجة الدقة المطلوبة.

- هناك الكثير من الطرق الإحصائية لتقدير حجم العينة سواء من المجتمعات الصغيرة المحدودة، أو من المجتمعات الكبيرة التي يصل تعدادها إلى الألوف أو الملايين، كما أن الطرق الإحصائية المستخدمة في تقدير حجم العينة ترتكز على معادلات ذات صيغ متعددة (بمعنى أن المعادلة الواحدة لها أكثر من صيغة) رغم أن أساسها الجبري واحد، ومن أهم هذه الطرق: طريقة الاحتمال المقدر، معادلة الرابطة الأمريكية للتربية، طريقة تقدير الانحراف المعياري.

- كثيرا ما يكون حجم المجتمع الذي ستسحب منه العينة معروفًا، كأن يكون عدد طلاب الجامعة، أو موظفي الحكومة أو موظفي شركات الاستثمار، أو سكان منطقة معينة.. في هذه الحالة يتم استخدام معامل تصحيح إضافي لتقليل حجم العينة بعض الشيء.
- تسهيلًا على الباحثين وضع الإحصائيون جداول لتحديد الحد الأدنى لحجم العينة المناسب بدرجة ثقة معينة إذا علمنا عدد المجتمع.
- من منظور العشوائية (Randomness)، هناك العينات العشوائية والعينات غير العشوائية.
- العينة العشوائية (Random Sample)، هي العينة التي تم اختيار مفرداتها دون تدخل من الباحث، بمعنى أن اختيارمفرداتها لا يتم بقصد أو تعمد أن تتضمن العينة مفردات معينة وعدم تضمنها مفردات أخرى.وفي العينة العشوائية تتاح فرصة متساوية لجميع المفردات لأن يتم اختيارها ضمن العينة، ولذلك تسمى بالعينة غير المتحيزة، أو العينة الاحتمالية (Probability Samples).
- من أهم أنماط العينة الاحتمالية (العشوائية) نذكر:العينة العشوائية البسيطة، العينة العشوائية المنتظمة، العينة العشوائية الطبقية، العينة العشوائية متعددة المراحل.
- العينات غير العشوائية أو غير الاحتمالية (Non-Probability Samples)، هي العينات التي لا تتاح في اختيارها فرص متساوية لجميع مفردات المجتمع أن يتم اختيارها ضمن العينة، وكمبدأ عام، فإن العينات غير الاحتمالية لا تعمم نتائجها المجتمع، ولا يمكن تقدير معالم المجتمع من إحصاءات تلك العينات، ومن أمثلة العينات غير العشوائية: العينة العارضة، والعينة المقصودة، وعينة الحصص.
- إن العينة مهما كانت دقة اختيارها عشوائيًا فإنها لا تكون ممثلة للمجتمع مائة بالمائة، ولذلك هناك ما يعرف بخطأ المعاينة (Sampling Error) والذي ينعكس على المقاييس الإحصائية التي تجرى على البيانات المأخوذة من العينة؛ بحيث تأتي مختلفة بعض الشيء عن المجتمع.

مصادر الفصل السادس ومراجعه

(أ) مصادر ومراجع عربية:

– عبد الحميد محمد نجم & محمد عبد الهادي المحميد (1990)، الإحصاء الوصفي والتحليلي مع استخدام البرامج الجاهزة، (الكويت: مكتبة جامعة الكويت).

– عزت عبد الحميد محمد حسن (2011)، الإحصاء النفسي والتربوي، (القاهرة: دار الفكر العربي).

(ب) مصادر ومراجع أجنبية:

- Alexander, Hershel Julius (1993) Multiple Matrix Sampling Rank-Order Estimation as a Function of Context Reliability, Standard Error, Sampling Design, and Sampling Percentage Dissertation Abstracts International, Vol. 54 (2-A), pp. 493.

- Alliger, George M.; Williams, Kevin J.(1993) Using Signal-Contingent Experience Sampling Methodology to Study Work in the Field: A Discussion and Illustration Examining Task Perceptions and Mood Personnel Psychology, Vol.. 46 (3), Special Issue: Innovations in Research Methods for Field Settings. pp. 525-549.

- Ary, Donald (1984) Mathematical Explanation of Error in Duration Recording Using Partial Interval, Whole Interval, and Momentary Time Sampling Behavioral Assessment, Vol. 6 (3), pp.221-228.

- Borror, Connie M. Ye, Nong; Mahwah (2003) Statistical Analysis of Normal and Abnormal Data. In: The Handbook of Data Mining. , NJ, US: Lawrence Erlbaum Associates, Publishers, pp. 67-102.

- Delespaul, Phillipe A. E. G. (1992) Technical Note: Devices and Time-Sampling Procedures In: The Experience of Psychopathology: Investigating Mental Disorders in Their Natural Settings. deVries, Marten W.; New York, : Cambridge University Press, pp. 363-373.

- Diaz, Joseph D. (2000) Religion and Gambling in Sin-City: A Statistical Analysis of the Relationship between Religion and Gambling Patterns in Las Vegas Residents; Social Science Journal, Vol. 37 (3), pp. 453-458.

- Ding, Mingzhou; Chen, Yanqing; Kelso, J. A. Scott (2002) Statistical Analysis of Timing Errors Brain and Cognition, Vol. 48 (1), Feb 2002. Special Issue: Human Movement Timing and Coordination. pp. 98-106.

- Edwards, Ron; Kearns, Kathleen; Tingstrom, Daniel H.(1991) Accuracy of Long Momentary Time-Sampling Intervals: Effects of Errors in the Timing of Observations.; Journal of Psycho-Educational Assessment, Vol. 9 (2),. pp.160-165.

- Gorard, Stephen (2006) Towards a Judgment-Based Statistical Analysis.; British Journal of Sociology of Education, Vol. 27(1), pp.67-80.

- Graham Kalton (1993); Sampling Considerations in Research on HIV Risk and Illness. in: Methodological Issues in AIDS Behavioral Research. Ostrow, David G.; Kessler, Ronald C.; New York, NY, US: Plenum Press, pp.53-74.

- Harrop, Alex; Daniels, Michael (1986) Methods of Time Sampling: A Reappraisal of Momentary Time Sampling and Partial Interval Recording. Journal of Applied Behavior Analysis, Vol. 19 (1), pp.73-77.

- Hormuth, Stefan E.; Brückner, Erika; Kölner (1985) Telephone Interviews in Sociology and Social Psychology: Problems of Sampling, Contacting, and Survey Design. Zeitschrift Für Soziologie and Sozialpsychologie, Vol. 37(3), Sep 1985. pp. 526-545.

- Jensun, Klaus Bruhn (2002) A Handbook of Media and Communication Research: Qualitative and Quantitative Methodologies. London. Rout Ledge.

- Kalton, Graham; Anderson, Dallas W.(1989) Sampling Rare Populations. In: Special Research Methods for Gerontology. Lawton, M. Powell; Herzog, A. Regula; Amityville, NY, US: Baywood Publishing Co, Inc, pp. 7-30

- Kenny, David A.; Mannetti, Lucia; Pierro, Antonio (2002). The Statistical Analysis of Data from Small Groups. Journal of Personality and Social Psychology, Vol. 83(1), pp. 126-137.

- Klesges, Robert C.; Woolfrey, Joan; Vollmer James (1985). An Evaluation of the Reliability of Time Sampling Versus Continuous Observation Data Collection Journal of Behavior Therapy and Experimental Psychiatry, Vol. 16(4), pp. 303-307.

- Krejcie, Robert V. and Daryle W. Morgan. (1970) “Determining Sample Size for Research Activities”. Educational and Psychological Measurements, VOL. 30 , p. 608.

- Kuzel, Anton J. (1992) Sampling in Qualitative Inquiry. In: Doing Qualitative Research. Crabtree, Benjamin F.; Miller, William L.; Thousand Oaks, CA, US: Sage Publications, Inc, pp. 31-44.

- Lavrakas, Paul J. (1993) Telephone Survey Methods: Sampling, Selection, and Supervision (2nd ed.). Thousand Oaks, CA, US: Sage Publications, Inc, 1993. ix, pp181.

- Lombard, Matthew; Snyder-Duch, Jennifer; Bracken, Cheryl Campanella (2003) Content Analysis In Mass Communication: Assessment and Reporting of Inter-Coder: Reliability Correction. Human Communication Research, VOL. 29 (3), pp. 469-472.

- Maher, Brendan A. (1992) Stimulus Sampling In Clinical Research: Representative Design Reviewed.; In: Kazdin, Alan E. Methodological Issues & Strategies In Clinical Research. Washington, DC, US: American Psychological Association, pp. 107-115.

- Marsh, Catherine; Scarbrough, Elinor; (1990) Testing Nine Hypotheses About Quota sampling. Journal of the Market Research Society, Vol. 32(4), pp. 485-506.

- Matthews, Timothy C. (1986) The Accuracy of Time Sampling Procedures for Estimating Behavioral Frequency Dissertation Abstracts International, Vol. 46(7-B),pp. 24 - 48.

- Mintz, Alexander (2005) The Feasibility of the Use of Samples in Content Analysis. In: George W. Stewart; Lasswell, Harold D.; Leites, Nathan. (eds), Language of Politics; Studies in Quantitative Semantics Oxford, England: pp. 127-152. (Psyc. INFO Database Record (c) 2005 APA.

- Morse, Janice M.(1991) Strategies for Sampling. In: Qualitative Nursing Research: A Contemporary Dialogue (Rev. ed.). Morse, Janice M.; Thousand Oaks, CA, US: Sage Publications, Inc, pp. 127-145.

- Palmore, Erdman B. (1989) Medical Records as Sampling Frames and Data Sources. In: Special Research Methods for Gerontology. Lawton, M. Powell; Herzog, A. Regula; Amityville, NY, US: Bay wood Publishing Co, Inc, pp. 127-161.

- Peregrine, Peter N.; Drews, David R.; North, Melissa (1993) Sampling Techniques and Sampling Error in Naturalistic Observation: An Empirical Evaluation with Implications for Cross-Cultural Research. Cross-Cultural Research: The Journal of Comparative Social Science, Vol. 27(3-4), pp. 232-246.

- Ray, J. J. (1985) Random Sampling Might Not Be Impossible After All. Political Psychology, Vol. 6(1), Mar 1985. pp. 141-146.

- Ray, John J. (1983) A Comparison Between Cluster And 'Random' Sampling. Journal of Social Psychology, Vol. 121(1), pp. 155-156

- Rojahn, Johannes; Kanoy, Robert C.(1985); Toward an Empirically Based Parameter Selection for Time-Sampling Observation Systems. Journal of Psychopathology and Behavioral Assessment, Vol. 7(2), pp. 99-120.

- Saudargas, Richard A.; Zanolli, Kathleen (1990) Momentary Time Sampling as an Estimate of Percentage Time: A Field Validation. Journal of Applied Behavior Analysis, Vol. 23(4), pp. 533-537.

- Seco, Guillermo Vallejo; Hermida, José Ramón Fernández; Villa, Roberto Secades (2004) Statistical Analysis and Considerations of Power in

Program Evaluation Using Two-Stage Sampling Designs Psychology in Spain, Vol. 8, pp. 77-88.

- Shepherd, Martin (1985) EMDISP: A Visual Display System with Digital and Analogue Sampling; Behavior Research Methods, Instruments & Computers, Vol. 16(3), pp. 297-302.

- Vallejo, G.; Livacic-Rojas, P.; Conejo, N. (2003) Statistical Analysis of AB Designs by Means of Pooled Time Series. International Journal of Psychology & Psychological Therapy, Vol. 3(1), pp. 7-22.

- Van Egeren, Lawrence F. (1973) Multivariate Statistical Analysis; Psychophysiology, Vol. 10(5), pp. 517-532.

- Van Meter, Karl M. (1990) Sampling and Cross-Classification Analysis in International Social Research; In: Comparative Methodology: Theory and Practice in International Social Research. Oyen, Else; Thousand Oaks, CA, US: Sage Publications, Inc, pp.172-186.

- Walker, Elaine F.; Lewine, Richard R. (1993) Sampling Biases in Studies of Gender and Schizophrenia. Schizophrenia Bulletin, Vol. 19(1), 1993. pp. 1-7.

- Watters, John K (.1993) The Significance of Sampling and Understanding Hidden Populations. Drugs & Society, Vol. 7(3-4), pp. 2-13.

- Weiss, Robin S.; Remington, Roger; Ellis, Stephen R. (1989) Sampling Distributions of the Entropy in Visual Scanning. Behavior Research Methods, Instruments & Computers, Vol. 21(3), pp. 348-352.

- Young, Raymond J.(1986) Methodology for Community Educational Study: How Appropriate is Cluster Sampling? Journal of Experimental Education, Vol. 54(2), pp. 114-117.

الـمـلاحـــق

(١) جدول القيم الحرجة لاختبار مربع كاي (كا^{٢})

(٢) جدول القيم الحرجة لمعامل ارتباط الرتب لسبيرمان

(٣) جدول القيم الحرجة لمعامل ارتباط بيرسون

(٤) جدول القيم الحرجة لاختبار (ت) t.test

(٥) جدول القيم الحرجة لاختبار توكي Q

(٦) جدول الدلالة الإحصائية للنسبة الفائية (ف) F

المصادر:

- عبد الحميد محمد نجم & محمد عبد الهادي المحميد (١٩٩٠)، الإحصاء الوصفي والتحليلي مع استخدام البرامج الجاهزة، (الكويت: مكتبة جامعة الكويت).
- عزت عبد الحميد محمد حسن (٢٠١١)، الإحصاء النفسي والتربوي، (القاهرة: دار الفكر العربي).

القيم الحرجة لاختبار مربع كاي (كا٢)

درجات الحرية	مستوى الدلالة 0.05	مستوى الدلالة 0.01	درجات الحرية	مستوى الدلالة 0.05	مستوى الدلالة 0.01
1	3.841	6.635	16	26.296	32.000
2	5.991	9.210	17	27.587	33.409
3	7.805	11.345	18	28.869	34.805
4	9.488	13.277	19	30.144	36.191
5	11.070	15.086	20	31.410	37.566
6	12.592	16.812	21	32.671	38.932
7	14.067	18.475	22	33.924	40.289
8	15.507	20.090	23	35.172	41.638
9	16.919	21.666	24	36.415	42.980
10	18.307	23.209	25	37.652	44.314
11	19.675	24.725	26	38.885	45.642
12	21.026	26.217	27	40.113	46.963
13	22.362	27.688	28	41.337	49.278
14	23.685	29.141	29	42.557	49.588
15	24.966	30.578	30	43.773	50.892

جدول رقم (٢)

القيم الحرجة لمعامل ارتباط الرتب لـ سبيرمان

عدد الأفراد (n)	مستوى الدلالة لاختبار ذي طرف واحد			
	0.05	0.025	0.01	0.005
	مستوى الدلالة لاختبار ذي طرفين			
	0.1	0.05	0.02	0.01
5	0.900	1.000	1.000	-
6	0.829	0.866	0.943	1.000
7	0.714	0.786	0.893	0.929
8	0.643	0.738	0.833	0.881
9	0.6	0.683	0.783	0.833
10	0.564	0.648	0.746	0.794
12	0.506	0.591	0.712	0.777
14	0.456	0.544	0.645	0.715
16	0.425	0.506	0.601	0.665
18	0.399	0.475	0.564	0.625
20	0.377	0.45	0.534	0.591
22	0.359	0.428	0.508	0.562
24	0.343	0.409	0.485	0.537
26	0.329	0.392	0.465	0.515
28	0.317	0.377	0.448	0.496
30	0.306	0.364	0.432	0.478

جدول رقم (٣)

القيم الحرجة لمعامل ارتباط بيرسون

درجات الحرية (n-2)	مستوى الدلالة للطرف الواحد		درجات الحرية (n-2)	مستوى الدلالة للطرف الواحد	
	0.025	0.005		0.025	0.005
	مستوى الدلالة للطرفين			مستوى الدلالة للطرفين	
	0.05	0.01		0.05	0.01
1	0.997	0.999	29	0.355	0.456
2	0.95	0.99	30	0.349	0.449
3	0.878	0.959	35	0.325	0.418
4	0.811	0.917	40	0.304	0.393
5	0.754	0.874	45	0.288	0.372
6	0.707	0.834	50	0.273	0.354
7	0.666	0.798	55	0.261	0.338
8	0.632	0.765	60	0.250	0.325
9	0.602	0.735	65	0.240	0.312
10	0.576	0.708	70	0.232	0.302
11	0.553	0.684	75	0.224	0.292
12	0.532	0.661	80	0.217	0.283
13	0.514	0.641	85	0.211	0.275
14	0.497	0.623	90	0.205	0.267
15	0.482	0.606	95	0.200	0.260
16	0.468	0.590	100	0.195	0.254
17	0.456	0.575	125	0.174	0.228
18	0.444	0.561	150	0.159	0.208
19	0.433	0.549	175	0.148	0.164
20	0.423	0.537	200	0.138	0.181
21	0.413	0.526	300	0.113	0.148
22	0.404	0.515	400	0.098	0.128
23	0.396	0.505	500	0.088	0.115
24	0.388	0.496	1000	0.062	0.081
25	0.381	0.487	2000	0.044	0.058
26	0.374	0.478	5000	0.028	0.036
27	0.367	0.470	10000	0.0196	0.0258
28	0.361	0.463			

جدول رقم (٤)

القيم الحرجة لاختبار (ت) T-Test

درجات الحرية	دلالة الطرفين 0.05 / دلالة الطرف الواحد 0.025	دلالة الطرفين 0.01 / دلالة الطرف الواحد 0.005	درجات الحرية	دلالة الطرفين 0.05 / دلالة الطرف الواحد 0.025	دلالة الطرفين 0.01 / دلالة الطرف الواحد 0.005
1	12.71	63.66	27	2.05	2.77
2	4.30	9.92	28	2.05	2.76
3	3.18	5.84	29	2.05	2.76
4	2.78	4.60	30	2.04	2.75
5	2.57	4.03	31	2.04	2.74
6	2.45	3.71	32	2.04	2.74
7	2.36	3.50	33	2.03	2.73
8	2.31	3.36	34	2.03	2.73
9	2.26	3.25	35	2.03	2.72
10	2.23	3.17	36	2.03	2.72
11	2.20	3.11	37	2.03	2.72
12	2.18	3.05	38	2.02	2.71
13	2.16	3.01	39	2.02	2.71
14	2.14	2.98	40	2.02	2.7
15	2.13	2.95	50	2.01	2.68
16	2.12	2.92	60	2.00	2.66
17	2.11	2.90	70	1.99	2.65
18	2.10	2.88	80	1.99	2.63
19	2.09	2.86	90	1.99	2.63
20	2.09	2.85	100	1.98	2.63
21	2.08	2.83	200	1.97	2.6
22	2.07	2.82	300	1.97	2.59
23	2.07	2.81	400	1.97	2.59
24	2.06	2.80	500	1.96	2.59
25	2.06	2.79	∞	1.96	2.58
26	2.06	2.78			

جدول رقم (٥)

القيم الحرجة لاختبار توكي Q

df denominator درجات حرية المقام	مستوى الدلالة	K (number of Means) ك (عدد المتوسطات)								
		2	3	4	5	6	7	8	9	10
1	0.05	18.00	27.00	32.80	37.10	40.40	43.10	45.40	47.40	49.10
	0.01	90.00	135.00	164.00	186.00	202.00	216.00	227.00	237.00	246.00
2	0.05	6.09	8.30	9.80	10.90	11.70	12.40	13.00	13.50	14.00
	0.01	14.00	19.00	22.30	24.70	26.60	28.20	29.50	30.70	31.70
3	0.05	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46
	0.01	8.26	10.60	12.20	13.30	14.20	15.00	15.60	16.20	16.70
4	0.05	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83
	0.01	6.51	8.12	9.17	9.96	10.60	11.10	11.50	11.90	12.30
5	0.05	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99
	0.01	5.70	6.97	7.80	8.42	8.91	9.32	9.67	9.97	10.20
6	0.05	3.46	4.34	4.90	5.31	5.63	5.89	6.12	6.32	6.49
	0.01	5.24	6.33	7.03	7.56	7.97	8.82	8.61	8.87	9.10
7	0.05	3.34	4.16	4.69	5.06	5.36	5.61	5.82	6.00	6.16
	0.01	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37
8	0.05	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92
	0.01	4.74	5.63	6.20	6.63	6.96	7.24	7.47	7.68	7.78
9	0.05	3.20	3.95	4.42	4.76	5.02	5.24	5.43	5.60	5.74
	0.01	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.32	7.49
10	0.05	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60
	0.01	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21
11	0.05	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49
	0.01	4.39	5.14	5.62	5.97	6.25	6.48	6.67	6.84	6.99
12	0.05	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.40
	0.01	4.32	5.04	5.50	5.84	6.10	6.32	6.51	6.67	6.81
13	0.05	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32
	0.01	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67
14	0.05	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25
	0.01	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54
16	0.05	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15
	0.01	4.13	4.78	5.19	5.49	5.72	5.92	6.08	6.22	6.35
18	0.05	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07
	0.01	4.07	4.70	5.06	5.38	5.60	5.79	5.94	6.08	6.20
20	0.05	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01
	0.01	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09
24	0.05	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92
	0.01	3.96	4.54	4.91	5.17	5.37	5.54	5.69	5.81	5.92
30	0.05	2.89	3.49	3.84	4.10	4.30	4.46	4.60	4.72	4.83
	0.01	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.56	5.76
40	0.05	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.74
	0.01	3.82	4.37	4.70	4.93	5.11	5.27	5.39	5.50	5.60
50	0.05	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65
	0.01	3.76	4.28	4.60	4.82	4.99	5.13	5.25	5.36	5.45
120	0.05	2.80	3.36	3.69	3.92	4.10	4.24	4.36	4.48	4.56
	0.01	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30
∞	0.05	2.77	3.31	3.63	3.36	4.03	4.17	4.29	4.39	4.47
	0.01	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16

جدول رقم (٦)

الدلالة الإحصائية للنسبة الفائية (ف) F

df denominator درجات حرية المقام (التباين الصغير)	مستوى الدلالة	df numerator درجات حرية البسط (التباين الكبير)								
		1	2	3	4	5	6	7	8	9
1	0.05	161.0	200.0	216.0	225.0	230.0	234.0	238.0	239.0	241.0
	0.01	4052.0	4999.0	5403.0	5625.0	5764.0	5859.0	5928. 0	5981.0	6022.0
2	0.05	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38
	0.01	98.49	99.01	99.17	99.25	99.30	99.33	99.34	99.36	99.38
3	0.05	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81
	0.01	34.12	30.81	29.46	28.71	28.24	27.91	27.67	27.49	27.34
4	0.05	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
	0.01	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	0.05	6.61	5.79	5.41	5.19	5.05	4.45	4.88	4.82	4.78
	0.01	16.26	13.27	12.06	11.39	10.97	10.67	10.45	10.27	10.15
6	0.05	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
	0.01	13.74	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	0.05	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
	0.01	12.25	9.55	8.45	7.85	7.46	7.19	7.00	6.84	6.71
8	0.05	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	0.01	11.26	8.65	7.59	7.01	6.63	6.37	6.19	6.03	5.91
9	0.05	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	0.01	10.56	8.02	6.99	6.42	6.06	5.80	5.62	5.47	5.35
10	0.05	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	0.01	10.04	7.56	6.55	5.99	5.64	5.39	5.21	5.06	4.95
11	0.05	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	0.01	9.65	7.20	6.22	5.67	5.32	5.07	4.88	4.74	4.63
12	0.05	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80
	0.01	9.33	6.93	5.95	5.41	5.06	4.82	4.65	4.50	4.39
13	0.05	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72
	0.01	9.07	6.70	5.74	5.20	4.86	4.62	4.44	4.30	4.19
14	0.05	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65
	0.01	8.86	6.51	5.56	5.03	4.69	4.46	4.28	4.14	4.03
15	0.05	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59
	0.01	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	0.05	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
	0.01	8.53	6.23	5.29	4.74	4.44	4.20	4.03	3.89	3.78
17	0.05	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50
	0.01	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	0.05	4.41	3.55	3.16	2.93	2.77	2.66	3.58	2.51	2.46
	0.01	8.28	6.01	5.09	4.58	4.25	4.01	3.85	3.71	3.60
19	0.05	4.38	3.52	3.13	2.90	2.74	2.63	2.55	2.48	2.43
	0.01	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	0.05	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.40
	0.01	8.10	5.85	4.94	4.43	4.10	3.87	3.71	3.56	3.45
21	0.05	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
	0.01	8.02	5.78	4.87	4.37	4.04	3.81	3.65	3.51	3.40
22	0.05	4.30	3.44	3.05	2.82	2.66	2.55	2.47	2.40	2.35
	0.01	7.94	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35

تابع جدول رقم (٦)

الدلالة الإحصائية للنسبة الفائية (ف) F

df denominator درجات حرية المقام (التباين الصغير)	مستوى الدلالة	df numerator درجات حرية البسط (التباين الكبير)								
		1	2	3	4	5	6	7	8	9
23	0.05	4.28	3.42	3.03	2.80	2.64	2.53	2.45	2.38	2.32
	0.01	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
24	0.05	4.26	3.40	3.01	2.78	2.62	2.51	2.43	2.36	2.30
	0.01	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.25
25	0.05	4.24	3.38	2.99	2.76	2.60	2.49	2.41	2.34	2.28
	0.01	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.21
26	0.05	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	0.01	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.17
27	0.05	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.30	2.25
	0.01	7.68	5.49	4.60	4.11	3.79	3.56	3.39	3.26	3.14
28	0.05	4.20	3.34	2.95	2.71	2.56	2.44	2.36	2.29	2.24
	0.01	7.64	5.45	4.57	4.07	3.76	3.53	3.36	3.23	3.11
29	0.05	4.18	3.33	3.93	2.70	2.54	2.43	2.35	2.28	2.22
	0.01	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.08
30	0.05	4.17	3.32	2.92	2.69	2.53	2.42	2.34	2.27	2.21
	0.01	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.06
32	0.05	4.15	3.30	2.90	2.67	2.51	2.40	2.32	2.25	2.19
	0.01	7.50	5.34	4.46	3.97	3.66	3.42	3.25	3.12	3.01
34	0.05	4.13	3.28	2.88	2.65	2.49	2.38	2.30	2.23	2.17
	0.01	7.44	5.29	2.42	3.93	3.61	3.38	3.21	3.08	2.97
36	0.05	4.11	3.26	2.86	2.63	2.48	2.36	2.28	2.21	2.15
	0.01	7.39	5.25	4.38	3.89	3.58	3.35	3.18	3.04	2.94
38	0.05	4.10	3.25	2.85	2.62	2.46	2.35	2.26	2.19	2.14
	0.01	7.35	5.21	4.34	3.86	3.54	3.32	3.15	3.02	2.91
40	0.05	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	0.01	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.88
42	0.05	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11
	0.01	7.27	5.15	4.29	3.80	3.49	3.26	3.10	2.96	2.86
44	0.05	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10
	0.01	7.24	5.12	4.26	3.78	3.46	3.24	3.07	2.94	2.84
46	0.05	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.14	2.09
	0.01	7.21	5.10	4.24	3.76	3.44	3.22	3.05	2.92	2.82
48	0.05	4.04	3.19	2.80	2.56	2.41	2.30	2.21	2.14	2.08
	0.01	7.19	5.08	4.22	3.74	3.42	3.20	3.04	2.90	2.80
50	0.05	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07
	0.01	7.17	5.06	4.20	3.72	3.41	3.18	3.02	2.88	2.78
55	0.05	4.02	3.17	2.78	2.54	2.38	2.27	2.18	2.11	2.05
	0.01	7.12	5.01	4.16	3.68	3.37	3.15	2.98	2.85	2.75
60	0.05	4.00	3.15	2.76	2.52	2.57	2.25	2.17	2.10	2.04
	0.01	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
65	0.05	3.99	3.14	2.75	2.51	2.36	2.24	2.15	2.08	2.02
	0.01	7.04	4.95	4.10	3.62	3.31	3.09	2.93	2.79	2.70
70	0.05	2.98	3.13	2.84	2.50	3.35	2.23	2.14	2.07	2.01
	0.01	7.01	4.92	4.08	3.60	3.29	3.07	2.91	2.77	2.67

تابع جدول رقم (٦)

الدلالة الإحصائية للنسبة الفائية (ف) F

df denominator درجات حرية المقام (التباين الصغير)	مستوى الدلالة	df numerator درجات حرية البسط (التباين الكبير)								
		1	2	3	4	5	6	7	8	9
80	0.05	3.96	3.11	2.72	2.48	2.33	2.21	2.12	2.05	1.99
	0.01	6.96	4.88	4.04	3.56	3.25	3.04	2.87	2.74	2.64
100	0.05	3.94	3.09	2.70	2.46	2.30	2.19	2.10	2.03	1.97
	0.01	6.90	4.82	3.98	3.51	3.20	2.99	2.82	2.69	2.59
125	0.05	3.92	3.07	2.68	2.44	2.29	2.17	2.08	2.01	1.95
	0.01	6.48	4.78	3.94	3.47	3.17	2.95	2.78	2.65	2.56
150	0.05	3.91	3.06	2.67	2.43	2.27	2.16	2.07	2.00	1.94
	0.01	6.81	4.75	3.91	3.44	3.14	2.92	2.76	2.62	2.53
200	0.05	3.89	3.04	2.65	2.41	2.26	2.14	2.05	1.98	1.92
	0.01	6.76	4.71	3.88	3.41	3.11	2.90	2.73	2.60	2.50
400	0.05	3.86	3.02	2.62	2.39	2.23	2.12	2.03	1.96	1.90
	0.01	6.70	4.66	3.83	3.36	3.06	2.85	2.69	2.55	2.46
1000	0.05	3.85	3.00	2.61	2.38	2.22	2.10	2.02	1.95	1.89
	0.01	6.66	4.62	3.80	3.34	3.04	2.82	2.66	2.53	1.43
∞	0.05	3.84	2.99	2.60	2.37	1.21	2.09	2.01	1.94	1.88
	0.01	6.64	4.60	3.78	3.32	3.02	2.80	2.64	2.51	2.41

تابع جدول رقم (٦)

الدلالة الإحصائية للنسبة الفائية (ف) F

df denominator درجات حرية المقام (التباين الصغير)	مستوى الدلالة	df numerator درجات حرية البسط (التباين الكبير)								
		10	11	12	14	16	20	24	30	40
1	0.05	242.00	243.00	244.00	245.00	246.00	248.00	249.00	250.00	251.00
	0.01	6056.0	6082.0	6106.0	6142.0	6169.0	6208.0	6234.0	6258.0	6286.0
2	0.05	19.39	19.40	19.41	19.42	19.43	19.44	19.45	19.46	19.47
	0.01	99.40	99.41	99.42	99.43	99.44	99.45	99.46	99.47	99.48
3	0.05	8.78	8.76	8.74	8.71	8.69	8.66	8.64	8.62	8.60
	0.01	27.23	27.13	27.05	26.92	26.83	26.69	26.60	26.50	26.41
4	0.05	5.96	5.93	5.91	5.87	5.84	5.80	5.77	5.74	5.71
	0.01	14.54	14.45	14.37	14.24	14.15	14.02	13.93	13.83	13.74
5	0.05	4.74	4.70	4.68	4.64	4.60	4.56	4.53	4.50	4.46
	0.01	10.05	9.96	9.89	9.77	9.68	9.55	9.47	9.38	9.29
6	0.05	4.06	4.03	4.00	3.96	3.92	3.87	3.84	3.81	3.77
	0.01	7.87	7.79	7.72	7.60	7.52	7.39	7.31	7.23	7.14
7	0.05	3.63	3.60	3.57	3.52	3.49	3.44	3.41	3.38	3.34
	0.01	6.62	6.54	6.47	6.35	6.27	6.55	6.07	5.98	5.90
8	0.05	3.34	3.31	3.28	3.23	3.20	3.15	3.12	3.08	3.05
	0.01	5.82	5.74	5.67	5.56	5.48	5.36	5.28	5.20	5.11
9	0.05	3.13	3.10	3.07	3.02	2.98	2.93	2.90	2.86	2.82
	0.01	5.26	5.18	5.11	5.00	4.92	4.80	4.73	4.64	4.56
10	0.05	2.97	2.94	2.91	2.86	2.82	2.77	2.74	2.70	2.67
	0.01	4.85	4.78	4.71	4.60	4.52	4.41	4.33	4.25	4.17

تابع جدول رقم (٦)

الدلالة الإحصائية للنسبة الفائية (ف) F

df denominator درجات حرية المقام (التباين الصغير)	مستوى الدلالة	df numerator درجات حرية البسط (التباين الكبير)								
		10	11	12	14	16	20	24	30	40
11	0.05	2.86	2.82	2.79	2.74	2.70	2.65	2.61	2.57	2.53
	0.01	4.54	4.46	4.40	4.29	4.21	4.10	4.02	3.94	3.86
12	0.05	2.76	2.72	2.69	2.64	2.60	2.54	2.50	2.46	2.42
	0.01	4.30	4.22	4.16	4.05	3.98	3.86	3.78	3.70	3.61
13	0.05	2.67	2.63	2.60	2.55	2.51	2.46	2.42	2.38	2.34
	0.01	4.10	4.02	3.96	3.85	3.78	3.67	3.59	3.51	3.42
14	0.05	2.10	2.56	2.53	2.48	2.44	2.39	2.35	2.31	2.27
	0.01	3.94	3.86	3.80	3.70	3.62	3.51	3.43	3.34	3.26
15	0.05	2.55	2.51	2.48	2.43	2.39	2.33	2.29	2.25	2.21
	0.01	3.80	3.73	3.67	3.56	3.48	3.36	3.29	3.20	3.12
16	0.05	2.49	2.45	2.42	2.37	2.33	2.28	2.24	2.20	2.16
	0.01	3.69	3.61	3.55	3.45	3.37	3.25	3.18	3.10	3.01
17	0.05	2.45	2.41	2.38	2.33	2.29	2.23	2.19	2.15	2.11
	0.01	3.59	3.52	3.45	3.35	3.27	3.16	3.08	3.00	2.92
18	0.05	2.41	2.37	2.34	2.29	2.25	2.19	2.15	2.11	2.07
	0.01	3.51	3.44	3.37	3.27	3.19	3.07	3.00	2.91	2.83
19	0.05	2.38	2.34	2.31	2.26	2.21	2.15	2.11	2.07	2.02
	0.01	3.43	3.36	3.30	3.19	3.12	3.00	2.92	2.84	2.76
20	0.05	2.35	2.31	2.28	2.23	2.18	2.12	2.08	2.04	1.99
	0.01	3.37	3.30	3.23	3.13	3.05	2.94	2.86	2.77	2.69
21	0.05	2.32	2.28	2.25	2.20	2.15	2.09	2.05	2.00	1.96
	0.01	3.31	3.24	3.17	3.07	2.99	2.88	2.80	2.72	2.63
22	0.05	2.30	2.26	2.23	2.18	2.13	2.07	2.03	1.98	1.93
	0.01	3.26	3.18	3.12	3.02	2.94	2.83	2.75	2.67	2.58
23	0.05	2.28	2.24	2.20	2.14	2.10	2.04	2.00	1.96	1.91
	0.01	3.21	3.14	3.07	2.97	2.89	2.78	2.70	2.62	2.53
24	0.05	2.26	2.22	2.18	2.13	2.09	2.02	1.98	1.94	1.89
	0.01	3.17	3.09	3.03	2.93	2.85	2.74	2.66	2.58	2.49
25	0.05	2.24	2.20	2.16	2.11	2.06	2.00	1.96	1.92	1.87
	0.01	3.13	3.05	2.99	2.89	2.81	2.70	2.62	2.54	2.45
26	0.05	2.22	3.18	2.15	2.10	2.05	1.99	1.95	1.90	1.85
	0.01	3.09	3.02	2.96	2.86	2.77	2.66	2.58	2.50	2.41
27	0.05	2.20	2.16	2.13	2.08	2.03	1.97	1.93	1.88	1.84
	0.01	3.06	2.98	2.93	2.83	2.74	2.63	2.55	2.47	2.38
28	0.05	2.19	2.15	2.12	2.06	2.02	1.96	1.91	1.87	1.81
	0.01	3.03	2.95	2.90	2.80	2.71	2.60	2.52	2.44	2.35
29	0.05	2.18	2.14	2.10	2.05	2.00	1.94	1.90	1.85	1.80
	0.01	3.00	2.92	2.87	2.77	2.68	2.57	2.49	2.41	2.32
30	0.05	2.16	2.12	2.09	2.04	1.99	1.93	1.89	1.84	1.79
	0.01	2.98	2.90	2.84	2.74	2.66	2.55	2.47	2.38	2.29
32	0.05	2.14	2.10	2.07	2.07	1.97	1.91	2.86	1.82	1.76
	0.01	2.94	2.86	2.80	2.70	2.62	2.51	2.42	2.34	2.25
34	0.05	2.12	2.08	2.05	2.00	1.95	1.89	1.84	1.80	1.74
	0.01	2.89	2.82	2.76	2.66	2.58	2.47	2.38	2.30	2.21
36	0.05	2.10	2.06	2.03	1.98	1.93	1.87	1.82	1.78	1.72

تابع جدول رقم (٦)

الدلالة الإحصائية للنسبة الفائية (ف) F

df denominator درجات حرية المقام (التباين الصغير)	مستوى الدلالة	df numerator درجات حرية البسط (التباين الكبير) 10	11	12	14	16	20	24	30	40
	0.01	2.86	2.78	2.72	2.62	2.54	2.43	2.35	2.26	2.17
38	0.05	2.09	2.05	2.02	1.96	1.92	1.85	1.80	1.76	1.71
	0.01	2.82	2.75	2.69	2.59	2.51	2.40	2.32	2.22	2.14
40	0.05	2.07	2.04	2.00	1.95	1.90	1.84	1.79	1.74	1.69
	0.01	2.80	2.73	2.66	2.56	2.49	2.37	2.29	2.20	2.11
42	0.05	2.06	2.02	1.99	1.94	1.89	1.82	1.78	1.73	1.68
	0.01	2.77	2.70	2.64	2.54	2.46	2.35	2.26	2.17	2.08
44	0.05	2.05	2.01	1.98	1.92	1.88	1.81	1.76	1.72	1.66
	0.01	2.75	2.68	2.62	2.52	2.44	2.32	2.24	2.15	2.06
46	0.05	2.04	2.00	1.97	1.91	1.87	1.80	1.75	1.71	1.65
	0.01	2.73	2.66	2.60	2.50	2.42	2.30	2.22	2.13	2.04
48	0.05	2.03	1.99	1.96	1.90	1.86	1.79	1.74	1.70	1.64
	0.01	2.71	2.64	2.58	2.48	2.40	2.28	2.20	2.11	2.02
50	0.05	2.02	1.98	1.95	1.90	1.85	1.78	1.74	1.69	1.63
	0.01	2.70	2.62	2.56	2.46	2.39	2.26	2.18	21.00	2.00
55	0.05	2.00	1.97	1.93	1.88	1.83	1.76	172.00	1.67	1.61
	0.01	2.66	2.59	2.53	2.43	2.35	2.23	2.15	2.06	1.96
60	0.05	1.99	1.95	1.92	1.86	1.81	1.75	1.70	1.65	1.59
	0.01	2.63	2.56	2.50	2.40	2.32	2.20	2.12	2.03	1.93
65	0.05	1.98	1.94	1.90	1.85	1.80	1.73	1.68	1.63	1.57
	0.01	2.61	2.54	2.47	2.37	2.30	2.18	2.09	2.00	1.90
70	0.05	1.97	1.93	1.89	1.84	1.79	1.72	1.67	1.62	1.56
	0.01	2.59	2.51	2.45	2.35	2.28	2.15	2.07	1.98	1.88
80	0.05	1.95	1.91	1.88	1.82	1.77	1.70	1.65	1.60	1.54
	0.01	2.55	2.48	2.41	2.32	2.24	2.11	2.03	1.94	1.84
100	0.05	1.92	1.88	1.85	1.79	1.75	1.68	1.63	1.57	1.51
	0.01	2.51	2.34	2.36	2.26	2.19	2.06	1.98	1.89	1.79
125	0.05	1.90	1.86	1.83	1.77	1.72	1.65	1.60	1.55	1.49
	0.01	2.47	2.40	2.23	2.23	2.15	2.03	1.94	1.85	1.75
150	0.05	1.89	1.85	1.82	1.76	1.71	1.64	1.59	1.54	1.47
	0.01	2.44	2.37	2.30	2.20	2.12	2.00	1.91	1.83	1.72
200	0.05	1.87	1.83	1.80	1.74	1.69	1.62	1.57	1.52	1.54
	0.01	2.41	2.34	2.28	2.17	2.09	1.97	1.88	1.79	1.69
400	0.05	1.85	1.81	1.78	1.72	1.67	1.60	1.54	1.49	1.42
	0.01	2.37	2.29	2.23	2.12	2.04	1.92	1.84	1.74	1.64
1000	0.05	1.84	1.80	1.76	1.70	1.65	1.58	1.53	1.47	1.41
	0.01	2.34	2.26	2.20	2.09	2.01	1.87	1.71	1.71	1.61
∞	0.05	1.83	1.79	1.75	1.69	1.64	1.57	1.52	1.46	1.40
	0.01	2.32	2.24	2.18	2.07	1.99	1.87	1.79	1.69	1.59

صدر من هذه السلسلة

الإعلان في الأنظمة الإذاعية المعاصرة	د. هويدا مصطفى
المرأة والإعلام في عالم متغير	د. ناهد رمزي
تكنولوجيا الاتصال (المخاطر والتحديات والتأثيرات الاجتماعية)	د. شريف درويش اللبان
تكنولوجيا النشر الصحفي (الاتجاهات الحديثة)	د. شريف درويش اللبان
مدخل إلى الإخراج الصحفي	د. سعيد غريب النجار
تكنولوجيا الصحافة في عصر التقنية الرقمية	د. سعيد غريب النجار
المسئولية الاجتماعية للصحافة	د. محمد حسام الدين
الإعلام والمجتمع	أ.د. منى سعيد الحديدي ، أ.د. سلوى إمام علي
نظريات في تشكيل اتجاهات الرأي العام	د .شيماء ذو الفقار
إدارة العلاقات العامة : المدخل الإستراتيجي	أ.د. راسم محمد الجمال، د. خيرت معوض عياد
نظام الاتصال والإعلام الدولي: الضبط والسيطرة	أ.د. راسم محمد الجمال
الإعلام ومعالجة الأزمات	د .حسن عماد مكاوي
الإعلان : أسسه .. وسائله .. فنونه	أ.د. منى سعيد الحديدي، أ.د. سلوى إمام علي
الإذاعة في القرن الحادي والعشرين	أ.د .حسن عماد مكاوي، د .عادل عبد الغفار

التسويق السياسي والإعلام : الإصلاح السياسي في مصر	أ.د .راسم محمد الجمال د. خيرت معوض عياد
الاتصال والإعلام في العالم العربي في عصر العولمة	أ.د .راسم محمد الجمال
الصحافة الإلكترونية.. دراسات في التفاعلية وتصميم المواقع	أ.د. شريف درويش اللبان
الفضائيات العربية ومتغيرات العصر : أعمال المؤتمر العلمي الأول للأكاديمية الدولية لعلوم الإعلام	أ.د. منى سعيد الحديدي
تحليل الخطاب الإعلامي : أطر نظرية ونماذج تطبيقية	د .محمد شومان
مستقبل طباعة الصحف العربية رقميًّا	د .مروة محمد كمال
الإعلام والمجتمع في عالم متغير	أ.د .حسن عماد مكاوي، د .عادل عبد الغفار
التليفزيون الفضائي العربي	د .هبة شاهين
التصوير الصحفي : الفيلمي والرقمي	د .سعيد غريب النجار
الإنترنت والصحافة الإلكترونية : رؤية مستقبلية	د .ماجد سالم تربان
الإعلام والمشاركة السياسية للمرأة : رؤية تحليلية واستشرافية	د .عادل عبد الغفار
صناعة الصحافة في العالم : تحديات الوضع الراهن وسيناريوهات المستقبل	د .محرز حسين غالي
أسس الدراما الإذاعية	أ.د .سامية أحمد علي
المادة الإخبارية في الراديو والتليفزيون	أ.د. بركات عبد العزيز
فن المونتاج التليفزيوني	د. نجلاء الجمال
كتابة السيناريو بين النظرية والتطبيق	محمد السيد عيد
الصحافة الورقية والإلكترونية في دول الخليج العربية النشأة والتطور	د. محمد يونس
مقدمة في التحليل الإحصائي (لبحوث الإعلام)	أ.د. بركات عبد العزيز